

GDSA Framework Development and Process Model Integration FY2021

Spent Fuel and Waste Disposition

***Prepared for
U.S. Department of Energy
Spent Fuel and Waste Science Technology***

***P.E. Mariner, T.M. Berg, K.W. Chang,
B.J. Debusschere, A.C. Eckert,
J. Harvey, T.C. LaForce, R.C. Leone,
M.M. Mills, M.A. Nole, H.D. Park,
F.V. Perry, D.T. Seidl, L.P. Swiler***

Sandia National Laboratories

September 13, 2021

M2SF-21SN010304053

SAND2021-12626 R



DISCLAIMER

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

DISCLAIMER

This is a technical report that does not take into account contractual limitations or obligations under the Standard Contract for Disposal of Spent Nuclear Fuel and/or High-Level Radioactive Waste (Standard Contract) (10 CFR Part 961). For example, under the provisions of the Standard Contract, spent nuclear fuel in multi-assembly canisters is not an acceptable waste form, absent a mutually agreed to contract amendment.

To the extent discussions or recommendations in this report conflict with the provisions of the Standard Contract, the Standard Contract governs the obligations of the parties, and this report in no manner supersedes, overrides, or amends the Standard Contract.

This report reflects technical work which could support future decision making by DOE. No inferences should be drawn from this report regarding future actions by DOE, which are limited both by the terms of the Standard Contract and Congressional appropriations for the Department to fulfill its obligations under the Nuclear Waste Policy Act including licensing and construction of a spent nuclear fuel repository.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.



**U.S. DEPARTMENT OF
ENERGY**



Sandia National Laboratories

APPENDIX E

NFCSC DOCUMENT COVER SHEET¹

Name/Title of Deliverable/Milestone/Revision No. GDSA Framework Development and Process Model Integration
FY2021 / M2SF-21SN010304053

Work Package Title and Number GDSA – Framework Development – SNL / SF-21SN01030405

Work Package WBS Number 1.08.01.03.04

Responsible Work Package Manager Paul Mariner
(Name/Signature)

Date Submitted

| | | | | |
|--|---|--------------------------------|---|---|
| Quality Rigor Level for Deliverable/Milestone ² | <input type="checkbox"/> QRL-1 <input type="checkbox"/> Nuclear Data | <input type="checkbox"/> QRL-2 | <input checked="" type="checkbox"/> QRL-3 | <input type="checkbox"/> QRL-4 Lab QA Program ³ |
|--|---|--------------------------------|---|---|

This deliverable was prepared in accordance with Sandia National Laboratories
(Participant/National Laboratory Name)

QA program which meets the requirements of
 DOE Order 414.1 NQA-1 Other

This Deliverable was subjected to:

Technical Review

Peer Review

Technical Review (TR)

Peer Review (PR)

Review Documentation Provided

Review Documentation Provided

- Signed TR Report or,
- Signed TR Concurrence Sheet or,
- Signature of TR Reviewer(s) below

- Signed PR Report or,
- Signed PR Concurrence Sheet or,
- Signature of PR Reviewer(s) below

Name and Signature of Reviewers

David Sassani

NOTE 1: Appendix E should be filled out and submitted with the deliverable. Or, if the PICS:NE system permits, completely enter all applicable information in the PICS:NE Deliverable Form. The requirement is to ensure that all applicable information is entered either in the PICS:NE system or by using the NFCSC Document Cover Sheet.

- In some cases there may be a milestone where an item is being fabricated, maintenance is being performed on a facility, or a document is being issued through a formal document control process where it specifically calls out a formal review of the document. In these cases, documentation (e.g., inspection report, maintenance request, work planning package documentation or the documented review of the issued document through the document control process) of the completion of the activity, along with the Document Cover Sheet, is sufficient to demonstrate achieving the milestone.

NOTE 2: If QRL 1, 2, or 3 is not assigned, then the QRL 4 box must be checked, and the work is understood to be performed using laboratory QA requirements. This includes any deliverable developed in conformance with the respective National Laboratory / Participant, DOE or NNSA-approved QA Program.

NOTE 3: If the lab has an NQA-1 program and the work to be conducted requires an NQA-1 program, then the QRL-1 box must be checked in the work Package and on the Appendix E cover sheet and the work must be performed in accordance with the Lab's NQA-1 program. The QRL-4 box should not be checked.

ACKNOWLEDGEMENTS

This report documents fiscal year (FY) 2021 accomplishments in GDSA Framework development and process model integration. Many of the accomplishments are described in more detail in other GDSA reports. For those accomplishments, highlights are selected and included in this report with citations. In addition, the following coauthors are acknowledged for specific contributions to this report:

- Section 3.1.3.2 Buffer Erosion and Waste Package Degradation: Rosie Leone, Paul Mariner, Melissa Mills, Michael Nole, and Heeho Park
- Section 3.1.3.3 Fortran Process Model: Jacob Harvey
- Section 3.2.1 GDSA Framework Graphical Workflow: Aubrey Eckert
- Section 3.2.2 Voronoi meshing and simulation: Tara LaForce
- Section 3.2.3 Uncertainty and Sensitivity Analysis: Laura Swiler
- Section 3.2.5 Geologic Framework Model: Frank Perry
- Appendix A. Fuel Matrix Degradation Modeling in FORTRAN: Jacob Harvey
- Appendix B. Surrogate Modeling of the Fuel Matrix Degradation (FMD) Process Model: Timothy Berg, K-Won Chang, Bert Debusschere, R.C. Leone, Paul Mariner, and Tom Seidl

The authors thank David Sassani of Sandia National Laboratories (SNL) for his thoughtful technical review and thank the staff from U.S. Department of Energy Office of Nuclear Energy (DOE-NE), Prasad Nair (DOE NE-81) and Tim Gunter (DOE NE-81), for their discussions, oversight, and guidance on topics addressed in this report.

EXECUTIVE SUMMARY

The Spent Fuel and Waste Science and Technology (SFWST) Campaign of the U.S. Department of Energy (DOE) Office of Nuclear Energy (NE), Office of Spent Fuel & Waste Disposition (SFWD) is conducting research and development (R&D) on geologic disposal of spent nuclear fuel (SNF) and high-level nuclear waste (HLW). A high priority for SFWST disposal R&D is disposal system modeling (DOE 2012, Table 6; Sevougian et al. 2019). The SFWST Geologic Disposal Safety Assessment (GDSA) work package is charged with developing a disposal system modeling and analysis capability for evaluating generic disposal system performance for nuclear waste in geologic media.

This report describes fiscal year (FY) 2021 advances of the Geologic Disposal Safety Assessment (GDSA) performance assessment (PA) development groups of the SFWST Campaign. The common mission of these groups is to develop a geologic disposal system modeling capability for nuclear waste that can be used to assess probabilistically the performance of generic disposal options and generic sites. The developing capability, called GDSA Framework, employs high-performance computing (HPC) capable codes PFLOTRAN and Dakota. The FY 2021 advances include:

- *Software and hardware infrastructure.* Full adoption and implementation of the Atlassian Jira issue and project tracking software (Section 3.1.1.1); important improvements and additions to the PFLOTRAN quality assurance (QA) test suite (Section 3.1.1.2); and increased utilization of new computing resources (Section 3.1.1.3)
- *Code performance improvements.* Characteristic curve smoothing to prevent PFLOTRAN convergence problems when simulating dry out (Section 3.1.2.1); and implementation of advanced linear and nonlinear solvers in PFLOTRAN to greatly speed up calculations (Section 3.1.2.2)
- *Process model development.* Sharpening of strategies for developing and implementing coupled process models in PFLOTRAN for PA (Section 3.1.3.1); development of an implementation plan for a buffer erosion and waste package corrosion model framework for GDSA (Section 3.1.3.2); continued development of a flexible, PFLOTRAN-compatible fuel matrix degradation (FMD) process model (Section 3.1.3.3; Appendix A); upgrades to the FMD surrogate models of the MATLAB FMD process model (Section 3.1.3.3; Appendix B); enhancements to the multi-continuum modeling of coupled fracture-matrix processes (Section 3.1.3.4); implementation of new thermal conductivity models and dependencies (Section 3.1.3.5); implementation and testing of a new smectite-to-illite transition reduced-order model (Section 3.1.3.6); addition of two new criticality-related models, a neutronics surrogate model and a spacer grid degradation model (Section 3.1.3.7); and development of a comprehensive biosphere model (Section 3.1.3.8)
- *GDSA Framework development.* Improvement and application of the new Next Generation Workflow (NGW) graphical interface for GDSA Framework simulations (Section 3.2.1); visualization and local grid refinement using VoroCrust (Section 3.2.2); uncertainty and sensitivity analysis applications and methods (Section 3.2.3); leadership in international generic performance assessment modeling for DECOVALEX-2023 (Section 3.2.4); development of a Geologic Framework Model for crystalline rock outcrops and overburden (Section 3.2.5); improved generic reference case models for shale, crystalline, salt, and unsaturated alluvium host rocks (Section 3.2.6); and establishment of the GDSA Calculation Archive (Section 3.2.7)

- *Outreach.* International engagement in DECOVALEX, geologic repository clubs, and generic application collaborations facilitated by the open source and freely available PFLOTRAN code (Sections 3.3.1, 3.3.2, and 3.3.3); execution of a virtual PFLOTRAN short course (Section 3.3.3); and improvements to the GDSA Framework website (Section 3.3.4)
- *Integration and planning.* Assessment of the status of inter-laboratory GDSA integration and alignment of GDSA activities with priorities (Section 4); and the update of the 5-year plan for disposal R&D (Section 5)

The advances in PA modeling capabilities allow for emulation of fuel matrix degradation processes at each waste package in a repository simulation, simulation of changes to buffer and disturbed rock zone (DRZ) properties over time, improved simulation of heat flow due to new temperature dependencies included in the calculation of thermal conductivity, improved PFLOTRAN convergence for multiphase systems and dry out, and more detailed assessment of total system performance. New methods allow for quantitative bulk characterization of flow and transport in regions affected by discrete fracture networks and for simulating the interaction of fracture and matrix processes. New modeling approaches include advanced meshing capabilities, a new plan for simulating buffer erosion and waste package corrosion, and development of a comprehensive biosphere model. Progress in simulation workflow, QA workflow, process model coupling workflow, and other forms of supporting infrastructure is expected to greatly facilitate future model development, user-friendliness, and user adoption.

An important responsibility of the GDSA team is to integrate with disposal R&D activities across the SFWST Campaign to ensure that R&D activities support the parts of the generic safety assessments being developed. In FY 2021, the GDSA team continued to participate with other scientists and engineers at LANL, LBNL, PNNL, ORNL, INL, ANL, DOE, and SNL in the development of new high-temperature model capabilities needed for possible direct disposal of dual-purpose canisters (DPCs), dfnWorks, Geologic Framework Models, FMD model integration, DECOVALEX-2023 Task F performance assessment, and advanced biosphere modeling.

Each year, GDSA Framework improves as additional modelers and programmers from around the world use, apply, and contribute to its development. GDSA Framework can be shared because the primary codes, PFLOTRAN and Dakota, are open source, available for free download, and have supporting documentation online. Outreach and collaborations support a primary objective of the GDSA work package by facilitating testing of, and feedback on, PFLOTRAN and GDSA Framework and by increasing the likelihood outside users will contribute directly to code development in the future.

The ability to simulate increasingly complex repository reference cases continues to affirm that HPC-capable codes can be used to simulate important multi-physics couplings directly in a total system safety assessment demonstration. Reference-case-repository applications show that PFLOTRAN and its coupled codes can simulate complex coupled processes in a multi-kilometer domain while simultaneously simulating sub-meter-scale coupled behavior in the vicinity of each modeled waste package. Continued development will further enhance the preparedness of GDSA Framework for application in the future when transitioning to a program with potential sites.

This report fulfills the GDSA Framework Development Work Package Level 2 Milestone – *GDSA Framework Development and Process Model Integration*, M2SF-21SN010304053.

CONTENTS

| | Page |
|---|-------------|
| Acknowledgements..... | iv |
| Executive Summary..... | v |
| Nomenclature..... | xiii |
| 1. Introduction..... | 1 |
| 2. GDSA Framework..... | 3 |
| 2.1 Conceptual Framework..... | 3 |
| 2.2 Computational Framework..... | 4 |
| 2.2.1 PFLOTRAN..... | 6 |
| 2.2.2 Dakota..... | 6 |
| 3. GDSA Framework Development..... | 7 |
| 3.1 PFLOTRAN Development..... | 7 |
| 3.1.1 Software and Hardware Infrastructure..... | 8 |
| 3.1.1.1 Code Management System..... | 8 |
| 3.1.1.2 Quality Assurance..... | 8 |
| 3.1.1.3 Computing Resources..... | 9 |
| 3.1.2 Code Performance Improvements..... | 9 |
| 3.1.2.1 Characteristic Curve Smoothing..... | 9 |
| 3.1.2.2 Linear and Nonlinear Solvers..... | 10 |
| 3.1.3 Process Model Development..... | 11 |
| 3.1.3.1 Strategy for Process Model Enhancements..... | 11 |
| 3.1.3.2 Buffer Erosion and Waste Package Degradation..... | 12 |
| 3.1.3.3 Fuel Matrix Degradation (FMD)..... | 16 |
| 3.1.3.4 Multi-Continuum Transport..... | 20 |
| 3.1.3.5 Thermal Conductivity Enhancements..... | 21 |
| 3.1.3.6 Smectite-to-Illite Transition..... | 22 |
| 3.1.3.7 Criticality-Related Models..... | 24 |
| 3.1.3.8 Biosphere..... | 28 |
| 3.2 Framework Development..... | 29 |
| 3.2.1 GDSA Framework Graphical Workflow..... | 29 |
| 3.2.2 Voronoi meshing and simulation..... | 33 |
| 3.2.3 Uncertainty and Sensitivity Analysis..... | 35 |
| 3.2.3.1 Multifidelity UQ Methods..... | 35 |
| 3.2.3.2 Discrete Fracture Network (DFN) Analysis..... | 36 |
| 3.2.3.3 Sensitivity Analysis of an Improved Crystalline Reference Case..... | 37 |
| 3.2.4 DECOVALEX-2023 Task F..... | 39 |
| 3.2.4.1 Crystalline..... | 39 |
| 3.2.4.2 Salt..... | 41 |
| 3.2.5 Geologic Framework Model..... | 43 |
| 3.2.5.1 Data to Support the Conceptual Model..... | 43 |
| 3.2.5.2 Conceptual Model of the Shallow Geosphere..... | 45 |
| 3.2.6 Repository Reference Cases..... | 46 |

| | | |
|---------|---|----|
| 3.2.6.1 | Shale..... | 47 |
| 3.2.6.2 | Crystalline | 48 |
| 3.2.6.3 | Salt | 49 |
| 3.2.6.4 | Unsaturated Alluvium..... | 49 |
| 3.2.7 | GDSA Calculation Archive | 52 |
| 3.3 | User Group | 53 |
| 3.3.1 | Open Source..... | 54 |
| 3.3.2 | International User Community..... | 54 |
| 3.3.3 | Short Courses | 55 |
| 3.3.4 | GDSA Framework Website | 55 |
| 3.3.5 | SFWST Document Archive | 56 |
| 4. | Disposal Research R&D Roadmap..... | 57 |
| 4.1 | Findings of the 2019 Roadmap Update for GDSA | 57 |
| 4.2 | Alignment of FY 2021 GDSA Efforts with 2019 Roadmap Activities | 58 |
| 5. | Five-Year Plan for Disposal Research R&D..... | 61 |
| 5.1 | 5-Year Plan | 61 |
| 5.2 | 5-Year Plan Update..... | 61 |
| 5.3 | Alignment of the Near-Term GDSA Thrusts with FY 2021 GDSA Activities and GDSA Roadmap Activities | 62 |
| 6. | Conclusions | 63 |
| 7. | References | 64 |

Appendix A. Fuel Matrix Degradation Modeling in FORTRAN

Appendix B. Surrogate Modeling of the Fuel Matrix Degradation (FMD) Process Model

Appendix C. Near-Term GDSA Thrusts in the Disposal Research R&D 5-Year Plan Update

Appendix D. Near-Term GDSA Thrusts and Their Mapped Roadmap Activities

FIGURES

| | | Page |
|-------------|--|-------------|
| Figure 2-1 | Schematic diagram of the conceptual model framework of a generic geologic disposal system..... | 4 |
| Figure 2-2 | GDSA Framework structure..... | 5 |
| Figure 3-1 | Jira summary table snapshot (Nole et al. 2021)..... | 8 |
| Figure 3-2 | Example Van Genuchten characteristic curves employing various extensions at the unsaturated limit (Nole et al. 2021). The purple curve is the original option where the maximum capillary pressure is truncated. The other extension options in the figure are defined using the following abbreviations: F=Flat or constant, E=Exponential, L=Linear, CPC=Capillary-Pressure Cap, NOC=NO Cap. | 10 |
| Figure 3-3 | Conceptual model of buffer erosion due to a flowing fracture (Fig 6-108, Posiva 2013)..... | 13 |
| Figure 3-4 | Montmorillonite release rates as a function of water velocity in the intersecting fracture (Posiva 2013, Figure 6-97)..... | 14 |
| Figure 3-5 | Fuel matrix degradation rates (a) and (c) and remaining fuel volumes (b) for breached 24 PWR (green and blue) and 37 PWR (yellow and orange) waste packages in a shale repository reference case using the ANN FMD surrogate (left) and traditional fractional dissolution rate (10^{-7} yr^{-1}) model (right)..... | 18 |
| Figure 3-6 | Concentration (M) of UO_2^{2+} as a function of time in a log-log plot. A long equilibration time (~ 100 yrs) and > 1000 yrs is needed to reach the boundary condition concentration. A dynamic time stepping routine was implemented to allow for long total simulation times..... | 19 |
| Figure 3-7 | Comparison of PFLOTRAN results to the analytical solution of Tang et al. (1981) along the fracture (top) and into the matrix (bottom) at 2 m along fracture..... | 21 |
| Figure 3-8 | Smectite to illite evolution at 0 years (a) and 10,000 years (b) of buffer placed 2 to 3 meters from the north end of the domain. Buffer 1 is originally 90% smectite and has a reaction threshold temperature of 40°C . Buffer 2 is originally 75% smectite and has a reaction threshold temperature of 60°C . The rest of the domain is water-saturated soil with a porosity of 25%, tortuosity of 50%, and permeability of 10^{-12} m^2 . Initially the domain has a hydrostatic pressure of 1 MPa and a temperature of 20°C . At the beginning of the simulation, the north face boundary condition is set to 5 MPa and 260°C | 23 |
| Figure 3-9 | Smectite to illite time histories showing thermal and permeability changes for Buffer #2..... | 24 |
| Figure 3-10 | Neutronics surrogate and spacer grid degradation test problem showing temperatures at 35 days (a) and 43 days (b) for a domain with three different waste forms with unique criticality events. At 35 days WF #1 has not reached criticality while WF #2 and #3 have. At 45 days criticality has ceased at WF #2 and #3 and begun at WF #1. | 26 |
| Figure 3-11 | Spacer grid vitality, temperature, and heat of criticality over time showing loss of spacer grid vitality at 43 and 36 days for WFs #1 and #2, respectively. | 27 |

| | | |
|-------------|---|----|
| Figure 3-12 | Schematic diagram of pathways to be included in new biosphere model being developed for GDSA Framework (graphic from Condon et al. 2020)..... | 28 |
| Figure 3-13 | GDSA biosphere model interaction matrix (graphic from Condon et al. 2020)..... | 29 |
| Figure 3-14 | Detailed notional view of the Crystalline Reference Case UA Nested Workflow structure showing the workflows along with important files, actions, and codes that are employed for each. | 31 |
| Figure 3-15 | NGW implementation of the Local Resource Deterministic Workflow element of the Crystalline Reference Case UA Nested Workflow..... | 32 |
| Figure 3-16 | Simulation meshes for four test structure volumes showing permeability for heterogeneous simulations. Top Left: Test case 1 with 4 horizontal layers. Top Right: Test Case 2 with a pinch-out. Bottom Left: Test Case 3 with an interior lens. Bottom Right: Test Case 4 with layering and an offset fault. Test Cases 1-3 are shown on the XZ plane on a slice through the plane y=10, while solution for Test Case 4 is shown on the YZ plane on a slice through the plane x=10 (after LaForce et al. (2021)). | 34 |
| Figure 3-17 | Liquid saturation at the top of the injection interval (plan view-i.e., x-y plane) at 5 years of injection. Green circles mark the well locations, which are 1 km apart. | 34 |
| Figure 3-18 | Comparison of Sobol' indices from the reference PCE computed with 828 high-fidelity samples (blue) and for the MF PCE (orange) | 36 |
| Figure 3-19 | Boxplot of Peak Total ¹²⁹ I based on 20 DFNs using a correlated constant transmissivity (left) vs. 20DFNs with a correlated depth dependent transmissivity (right)..... | 37 |
| Figure 3-20 | Global sensitivity analysis results over time for the maximum ¹²⁹ I concentration in the aquifer..... | 38 |
| Figure 3-21 | Uniform ECPM representation of the four deterministic fractures of 4-fracture DFN problem..... | 40 |
| Figure 3-22 | Breakthrough curves of SNL simulations of the 4-fracture problem without stochastic fractures. Three simulations use PFLOTRAN and the advection-dispersion equation (ADE). Two use the DFN directly while two use an ECPM generated from DFN mapping. The ECPM Octree mesh (30,087 active cells) is much more refined at fracture locations than the ECPM Uniform mesh (9704 active cells)..... | 40 |
| Figure 3-23 | 4-fracture DFN with stochastic fractures..... | 41 |
| Figure 3-24 | Geologic cross-section of salt reference case for Task F | 42 |
| Figure 3-25 | Repository layout for generic salt reference case for Task F | 42 |
| Figure 3-26 | Surficial geology of the Sebago Lake Pluton reference area..... | 44 |
| Figure 3-27 | Example of groundwater flow (icicles) in a horizontal fracture in a shallow crystalline rock exposure in southern Maine (MGS 2017)..... | 44 |
| Figure 3-28 | Conceptual model of stratigraphy and fracture systems of the shallow geosphere. Vertical scale is highly exaggerated | 46 |
| Figure 3-29 | Shale repository submodel (A) and mesh (B) | 48 |

| | | |
|-------------|---|----|
| Figure 3-30 | Crystalline reference case ECPM realization showing the repository, a midline cutaway of the cells propagating fracture properties, and the surfaces where water flow QoI are measured | 49 |
| Figure 3-31 | Half-symmetry model domain of the unsaturated alluvium reference case (3915 m × 1065 m × 1005 m). The repository is located at a depth of 250 m. The half-symmetry repository includes 27 drifts containing 25 WPs (24-PWR, 100 y OoR, 40 GWd/MTU burn-up) spaced 20 m apart center-to-center. Drift spacing is 50 m..... | 50 |
| Figure 3-32 | Temperature, gas saturation, gas pressure and liquid saturation time histories at the “Fwp_inside” location (centermost WP) in the repository | 51 |
| Figure 3-33 | Gas pressure (A, B) and gas saturation (C, D) results from PFLOTRAN models using unstructured grids and Voronoi meshes at t = 100 years for an unsaturated alluvium repository..... | 52 |
| Figure 3-34 | GDSA Calculation Archive gitlab home page | 53 |
| Figure 3-35 | User count on the PFLOTRAN website around the world between June 12, 2020 and June 12, 2021, colored by country | 55 |
| Figure 3-36 | GDSA Framework website (http://pa.sandia.gov/) | 56 |

TABLES

| | | Page |
|-----------|--|-------------|
| Table 3-1 | Computation time with Newton-Raphson, Easy case, 32 cores | 11 |
| Table 3-2 | Computation time with FGMRES-ABF, Hard case, 32 cores..... | 11 |
| Table 3-3 | Values of the error metrics computed on the test set (and training set for the ANN) for the UO ₂ surface flux in the “natural” (i.e., non-log ₁₀ transformed) space. The error on the training data is zero for kNNr because for inverse-distance weighted averaging, the table prediction is the same as the tabulated value at the query point. | 17 |
| Table 3-4 | Representative values of properties for glacial deposits included in the conceptual model | 46 |
| Table 3-5 | Repository concepts and generic reference cases implemented with GDSA Framework..... | 47 |
| Table 4-1 | GDSA PA activities in the Roadmap Database of the 2019 Disposal R&D Roadmap update | 58 |
| Table 4-2 | FY 2021 GDSA activities mapped to GDSA Roadmap activities and 5-Year Plan near-term GDSA thrusts | 60 |

NOMENCLATURE

| | |
|----------------|--|
| 1D, 2D, 3D, 4D | one-, two-, three-, and four-dimensional |
| ABF | alternate block factorization |
| ADE | advection-dispersion equation |
| ANL | Argonne National Laboratory |
| ANN | artificial neural network |
| BCGS-ILU | BiCGSTAB for ILU preconditioner |
| BiCGSTAB | biconjugate gradient stabilized method |
| CPR | constrained pressure residual |
| DCDM | Dual Continuum Disconnected Matrix |
| DECOVALEX | Development of COupled models and their VALidation against Experiments |
| DFN | discrete fracture network |
| DOE | U.S. Department of Energy |
| DPC | dual-purpose canister |
| DR | disposal research |
| DRZ | disturbed rock zone |
| EBS | engineered barrier system |
| ECPM | equivalent continuous porous medium |
| Eq. | equation |
| FCA | FGMRES for CPR-ABF |
| FCQ | FGMRES for CPR-Quasi-IMPES |
| FEP | features, events, and processes |
| FGMRES | flexible generalized minimal residual method |
| FMD | Fuel Matrix Degradation |
| FY | fiscal year |
| g | gram |
| GB | gigabytes |
| GDSA | Geologic Disposal Safety Assessment |
| gpm | gallons per minute |
| GUI | graphical user interface |
| GWd | gigawatt day |
| HDF5 | hierarchical data format, version 5 |
| HLW | high-level nuclear waste |
| HPC | high-performance computing |
| IAEA | International Atomic Energy Agency |
| ID | identification number |
| ILU | incomplete lower-upper factorization |
| IMPES | implicit pressure and explicit saturation |
| IRF | instant release fraction |
| IWM | Integrated Waste Management |
| K | Kelvin |
| K_D | distribution coefficient |
| km | kilometer |
| kNNr | k Nearest-Neighbors regressor |

| | |
|----------|---|
| LANL | Los Alamos National Laboratory |
| LBNL | Lawrence Berkeley National Laboratory |
| LGPL | lesser general public license |
| LHS | Latin hypercube sampling |
| m | meter |
| M | molar |
| MAE | mean absolute error |
| MAPE | mean absolute percentage error |
| MATLAB | programming language by MathWorks |
| MB | megabytes |
| mm | millimeter |
| mM | millimolar |
| mol | mole |
| MPa | megapascal |
| MSE | mean square error |
| MTHM | metric tons of heavy metal |
| MWd | megawatt day |
| NA | not applicable |
| NBS | natural barrier system |
| NE | Office of Nuclear Energy |
| NEA | Nuclear Energy Agency |
| NGW | Next-Generation Workflow |
| NM | New Mexico |
| NTRDC | Newton Trust Region Dogleg-Cauchy |
| NTRDC-AS | NTRDC with auto-scaling |
| OECD | Organization for Economic Co-operation and Development |
| OFCT | Office of Fuel Cycle Technology |
| OoR | out of reactor |
| ORNL | Oak Ridge National Laboratory |
| OWL | Online Waste Library |
| PA | performance assessment |
| PCE | Polynomial Chaos Expansion |
| PDE | partial differential equation |
| PETSc | Portable Extensible Toolkit for Scientific Computation |
| PFLOTRAN | massively parallel reactive flow and transport model for describing subsurface processes (pflotran.org) |
| pH | negative logarithm of hydrogen ion activity |
| PNNL | Pacific Northwest National Laboratory |
| PWR | pressurized water reactor |
| QA | quality assurance |
| QoI | quantities of interest |
| RAM | random access memory |
| R&D | research and development |
| s | seconds |
| S&T | Storage and Transportation |
| SA | sensitivity analysis |
| SDA | SFWD Document Archive |

| | |
|-----------------|---|
| SFWD | Spent Fuel and Waste Disposition |
| SFWST | Spent Fuel and Waste Science and Technology |
| SKB | Swedish Nuclear Fuel and Waste Management Company |
| SNF | spent nuclear fuel |
| SNL | Sandia National Laboratories |
| TBD | to be determined |
| TCC | thermal characteristic curve |
| THMC | thermal-hydrologic-mechanical-chemical |
| TRL | technical readiness level |
| UA | uncertainty analysis |
| UFD | Used Fuel Disposition |
| ULR | unclassified limited release |
| UNF | used nuclear fuel |
| UO ₂ | uranium dioxide |
| UQ | uncertainty quantification |
| URL | underground research laboratory |
| U.S. | United States of America |
| USA | United States of America |
| V&V | verification and validation |
| W | watt |
| WF | waste form |
| WIPP | Waste Isolation Pilot Plant |
| WP | waste package |
| yr | year |

(This page left intentionally blank.)

1. INTRODUCTION

The Spent Fuel and Waste Science and Technology (SFWST) Campaign of the U.S. Department of Energy (DOE) Office of Nuclear Energy (NE), Office of Spent Fuel & Waste Disposition (SFWD) is conducting research and development (R&D) on geologic disposal of spent nuclear fuel (SNF) and high-level nuclear waste (HLW). A high priority for SFWST disposal R&D is disposal system modeling (DOE 2012, Table 6; Sevougian et al. 2019). The SFWST Geologic Disposal Safety Assessment (GDSA) work package is charged with developing a disposal system modeling and analysis capability for evaluating generic disposal system performance for nuclear waste in geologic media.

The capability being developed is a software package referred to as GDSA Framework. The primary codes used by GDSA Framework are PFLOTRAN and Dakota (Section 2.2). Each code is designed for massively-parallel processing in a high-performance computing (HPC) environment.

The purpose of the GDSA Framework Development work package is to develop a GDSA capability that:

- Integrates updated conceptual models of subsystem processes and couplings developed under this and other disposal research work packages;
- Is used to evaluate disposal research R&D priorities;
- Leverages existing computational capabilities (e.g., meshing, visualization, HPC) where appropriate; and
- Is developed and distributed in an open-source environment.

The long-term goal for the GDSA team is to develop a safety assessment capability that can simulate all potentially important features, events, and processes (FEPs) for a number of generic repository concepts/environments. Such a capability is years away, but a DOE timeline suggests that a performance assessment (PA) model for a potential candidate site will likely not be needed before at least 2037 (DOE 2013). Although the specific timing is more uncertain, continued advances over time in computing speed, and continued code development, are expected to be made in the capability to have it ready for when it will be applied for its ultimate purpose in a different future program.

For the near term, GDSA objectives are focused on including essential FEPs in the PA model and on developing a suite of probabilistic generic repository reference case applications. These near-term objectives are in line with the long-term goal. In addition, the products of the near-term objectives are useful for evaluating the effects of FEPs and input parameters on repository performance, which is useful for R&D planning.

For FY 2021, seven tasks were addressed:

- Use the 5-year Disposal Research R&D Plan (Sassani et al. 2020) to help identify additional capabilities needed in the short term and long term to advance GDSA Framework (e.g., multiphase processes, temperature dependencies, colloids, engineered barrier system (EBS) degradation processes, computational efficiency, gridding capability). The GDSA Framework work package coordinated with other work packages as applicable in identifying those needs, determining what is required to sufficiently address them, and working to fulfill them.
- Integrate subsystem models developed under this and other SFWST work packages into GDSA software and safety assessments (e.g., waste form (WF) degradation, waste package (WP) degradation, colloid stability and transport, EBS chemistry, EBS flow and transport, fracture

representation, thermal-hydrologic-mechanical-chemical (THMC) processes, natural system flow and transport).

- Develop and implement methods for computationally efficient multi-scale, multi-physics modeling (e.g., surrogate models, reduced-order models, physics-based machine learning, nested models, Voronoi cell refinement). This task aims to improve integration of complicated FEPs in probabilistic safety assessments.
- Develop, perform, and document verification and validation analyses of relevant GDSA model processes, expand regression testing to demonstrate and assure continued quality, and work to ensure a GDSA Framework release package that is traceable, reproducible, and user-friendly.
- Demonstrate the freely-available PFLOTRAN GDSA Framework and modeling capability at national and international forums and support an international DECOVALEX proposal for a multi-year PA modeling comparison of reference repository systems. Plan to conduct one or more workshops to promote accelerated use of the capability worldwide as expanding the user base provides additional testing of the code and opportunities for additional development by outside contributors.
- Attend, as appropriate, technical training (e.g., classes/workshops in Python, simulation and analysis software, or computational and analysis methods), technical conferences, and international clubs and initiatives with direct benefit to GDSA.
- Update the 5-year Disposal Research R&D plan to reflect developments in DOE priorities.

As documented in this report, good progress was made on each of these tasks. Section 2 describes the conceptual model framework and the PFLOTRAN-based computational framework for GDSA Framework. Section 3 describes FY 2021 advances in general code development, meshing, uncertainty quantification (UQ), workflow, and infrastructure. Section 4 inspects how the FY 2021 GDSA efforts address objectives highlighted for GDSA PA in the 2109 Roadmap Update (Sevougian et al. 2019). Section 5 summarizes the development of the new updated Disposal Research R&D 5-Year Plan and identifies the thrusts for GDSA and how they align with current GDSA efforts. Conclusions are summarized in Section 6.

This report fulfills the requirements of the GDSA Framework Development work package (SF-21SN01030405) Level 2 Milestone – *GDSA Framework Development and Process Model Integration FY2021*, M2SF-21SN010304053. The work presented herein builds on previous reports, i.e., Freeze et al. (2013); Sevougian et al. (2013); Sevougian et al. (2014); Mariner et al. (2015); Mariner et al. (2016); Mariner et al. (2017); Mariner et al. (2018); Mariner et al. (2019); Sevougian et al. (2019); Sevougian et al. (2019); Mariner et al. (2020a); and Mariner et al. (2020b).

2. GDSA FRAMEWORK

A PA for underground disposal of nuclear waste utilizes a comprehensive analysis of FEPs potentially affecting the release and transport of radionuclides to the biosphere. In a PA, plausible scenarios and processes that may affect repository performance are addressed. FEPs and scenarios are evaluated and screened for inclusion in the quantitative system analysis. Potentially pertinent FEPs are identified for inclusion/simulation in a quantitative PA model. Probabilistic simulations are performed, and results are evaluated against system performance metrics. Uncertainty and sensitivity analyses may also be performed to inform prioritization of additional research and/or model development.

The PA framework consists of a conceptual model framework and a computational framework. These frameworks are summarized in Sections 2.1 and 2.2, respectively. An overview of PA methodology and terminology is presented in Sevougian et al. (2014, Section 2.2), Meacham et al. (2011, Section 1) and elsewhere (Rechard 2002).

2.1 Conceptual Framework

A safety case for a deep geologic disposal facility is a comprehensive set of bases and analyses designed, in part, to assess regulatory compliance with respect to safety standards. More specifically, it is a widely accepted approach for documenting the bases for the understanding of the disposal system, describing the key justifications for its safety, and acknowledging the unresolved uncertainties and their safety significance (OECD 2004; IAEA 2006; Freeze et al. 2013). A full safety case may only be constructed for a specific site with an integrated design, but aspects of a safety case may also be developed for generic systems evaluated in the SFWST Campaign. In general, building these aspects for a generic safety case focuses on three primary components related most directly to post-closure safety assessments: a safety strategy, technical bases, and a safety assessment.

- The safety strategy provides direction and boundaries for the safety case. It guides the safety case by identifying requirements for site location, repository design, and safety objectives.
- Technical bases are the laws of nature and the physical and chemical barriers that govern the system. They address each FEP that could potentially facilitate or inhibit the transport of radionuclides from the repository to the biosphere. Development of the technical bases involves site characterization (mainly as defined/constrained characteristics for generic sites), FEPs identification including waste inventory, barriers to radionuclide release and migration, radionuclide behavior, and using natural analogs, model validation, code verification, and UQ.
- Safety assessment involves the analysis of technical bases to evaluate whether the objectives of the safety strategy are met. In safety assessment, each FEP screened in the technical bases is either included/incorporated into the probabilistic PA model or is excluded and addressed in separate analyses or process model simulations. In the PA model, regulatory metrics (e.g., annual dose rate) are estimated with probabilistic calculations to compare to regulatory limits.

The goals and objectives of the GDSA team focus on safety assessment and, more specifically, on the development of the PA model. Conceptually, the long-term vision for the GDSA effort is to ensure that the GDSA modeling capability can adapt to, and take advantage of, future advances in computational software and hardware and future advances in process modeling. In line with this vision, the near-term mission is to develop a robust suite of fully functional generic repository reference case applications (1) for evaluation of the effects of FEPs and input parameters on repository performance to inform R&D planning and (2) for application to future candidate sites by the time they are selected.

Consistent with the long-term vision, two open-source, HPC codes serve as the core of the GDSA Framework: PFLOTRAN and Dakota. PFLOTRAN is a thermal-hydrologic-chemical flow and transport code, and Dakota is a versatile probabilistic code (Section 2.2). The PFLOTRAN code is being developed by the GDSA team to accommodate new geologic disposal process models and capabilities through additional code development and coupling with external process models. The HPC capabilities of PFLOTRAN and Dakota allow for ever higher fidelity in total system performance assessment modeling as more powerful HPC resources become available.

As the GDSA modeling capability evolves, the GDSA team will continue to generate and refine three-dimensional models of disposal repository concepts complete with surrounding geospheres and connected biospheres. Sensitivity analyses will be performed on these models to distinguish the importance of features, processes, and parameters on model results. These analyses are expected to assist prioritization of future disposal R&D.

A conceptual model framework requires a coherent representation of pertinent FEPs. Figure 2-1 schematically illustrates the conceptual model framework for a repository system. To calculate a dose to a receptor in the biosphere, radionuclides released from the WF must pass through the repository EBS and the surrounding natural barrier system (NBS).

A FEPs database like the one developed and described in Freeze et al. (2011) can be used to help identify a full set of potentially important FEPs for a specific conceptual repository model. Many of the FEPs in a FEPs database may be directly simulated in the PA model. In a comprehensive PA, excluded FEPs (i.e., FEPs not simulated in the PA model) must be addressed in separate analyses and arguments.

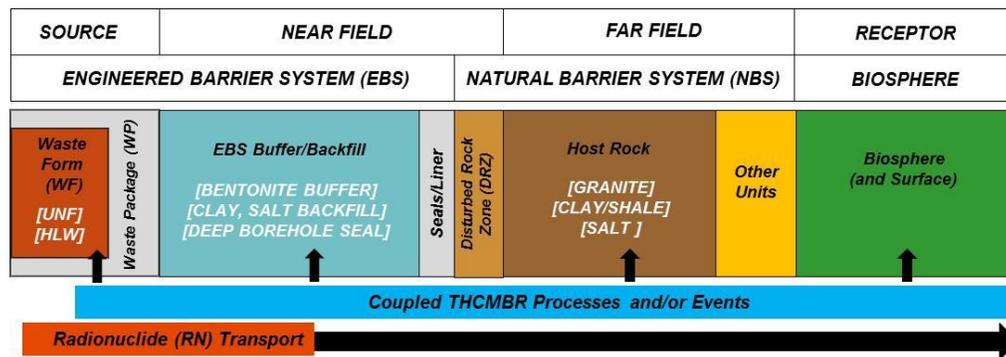


Figure 2-1 Schematic diagram of the conceptual model framework of a generic geologic disposal system

2.2 Computational Framework

Performance assessment of a geologic repository is aided by directly modeling the important coupled processes in the system and executing multiple probabilistic realizations. The approach of propagating uncertainty in detailed PA models is a continuation of the successful modeling approaches adopted for the Waste Isolation Pilot Plant (WIPP) PAs (Rechard 1995; Rechard 2002; Rechard and Tierney 2005) and for disposal of SNF and HLW in volcanic tuff (Rechard and Stockman 2014). Due to the complexity of the coupled processes, the large range of scales of these processes, and the need for large numbers of realizations, GDSA Framework is designed for massively-parallel processing in a HPC environment.

GDSA Framework consists of the following components:

- Input parameter database

- Software for sampling, sensitivity analysis, UQ, workflow, and traceability (Dakota)
- Petascale multiphase flow and reactive transport code (PFLOTRAN), working in concert with coupled process model codes (e.g., Fuel Matrix Degradation (FMD) Model)
- Computational support software and scripts for meshing, processing, and visualizing results (e.g., CUBIT, VoroCrust, dfnWorks, Python, ParaView, VisIt).

The two primary components of this computational framework are PFLOTRAN and Dakota. PFLOTRAN is a thermal-hydrologic-chemical multi-physics code used to simulate coupled multi-physics processes affecting waste isolation in a repository system and transport of released radionuclides to the biosphere over time. Simulated processes include heat flow, fluid flow, waste dissolution, radionuclide release, radionuclide decay and ingrowth, precipitation and dissolution of secondary phases, and radionuclide transport. Dakota is an uncertainty sampling and propagation code. Dakota is used to propagate uncertainty in PFLOTRAN simulations and to analyze PFLOTRAN results to assess sensitivities of model processes and inputs. Dakota is also used to graphically run and document the entire workflow of probabilistic simulations. These two codes are described in more detail in Sections 2.2.1 and 2.2.2.

The flow of data and calculations through the components of GDSA Framework is illustrated in Figure 2-2. In a probabilistic simulation, Dakota’s Next Gen Workflow manages the entire simulation from the generation of stochastic input for each PA realization to the execution of PFLOTRAN and production of custom output files via Python scripts. The sampled inputs are used by PFLOTRAN and its coupled process models to simulate source term release, EBS evolution, flow and transport through the EBS and NBS, and uptake in the biosphere. After the simulation, various software (e.g., Python, Matplotlib, ParaView) may be used to reduce and illustrate the output results of parameters and performance metrics. Dakota may also be used to evaluate the effects of parameter uncertainty on specific outputs.

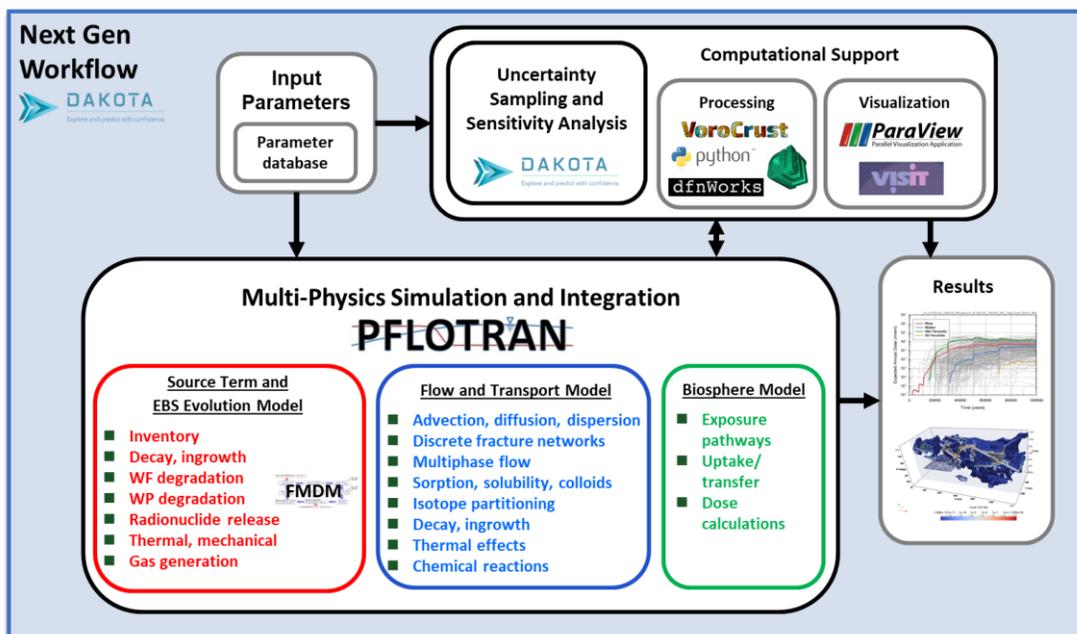


Figure 2-2 GDSA Framework structure

2.2.1 PFLOTRAN

PFLOTRAN (Hammond et al. 2011; Lichtner and Hammond 2012) is an open source, reactive multi-phase flow and transport simulator designed to leverage massively-parallel HPC to simulate subsurface earth system processes. PFLOTRAN has been employed on petascale leadership-class DOE computing resources (e.g., Jaguar [at Oak Ridge National Laboratory (ORNL)] and Franklin/Hopper [at Lawrence Berkeley National Laboratory (LBNL)]) to simulate THC processes at the Nevada Test Site (Mills et al. 2007), multi-phase CO₂-H₂O for carbon sequestration (Lu and Lichtner 2007), CO₂ leakage within shallow aquifers (Navarre-Sitchler et al. 2013), and uranium fate and transport at the Hanford 300 Area (Hammond et al. 2007; Hammond et al. 2008; Hammond and Lichtner 2010; Hammond et al. 2011; Chen et al. 2012; Chen et al. 2013). PFLOTRAN is also under development for use in PA at the Waste Isolation Pilot Plant (WIPP).

PFLOTRAN solves the non-linear partial differential equations describing non-isothermal multi-phase flow, reactive transport, and geomechanics in porous media. Parallelization is achieved through domain decomposition using the Portable Extensible Toolkit for Scientific Computation (PETSc) (Balay et al. 2013). PETSc provides a flexible interface to data structures and solvers that facilitate the use of parallel computing. PFLOTRAN is written in Fortran 2003/2008 and leverages state of the art Fortran programming (i.e. Fortran classes, pointers to procedures, etc.) to support its object-oriented design. The code provides “factories” within which the developer can integrate a custom set of process models and time integrators for simulating surface and subsurface multi-physics processes. PFLOTRAN employs a single, unified framework for simulating multi-physics processes on both structured and unstructured grid discretizations (i.e., there is no duplication of the code that calculates multi-physics process model functions in support of structured and unstructured discretizations). The code requires a small, select set of third-party libraries (e.g., MPI, PETSc, BLAS/LAPACK, HDF5, Metis/Parmetis). Both the unified structured/unstructured framework and the limited number of third-party libraries greatly facilitate usability for the end user.

2.2.2 Dakota

The Dakota software toolkit is open source software developed and supported at Sandia National Laboratories (Adams et al. 2012; Adams et al. 2013). Dakota provides deterministic codes an extensible interface for propagating uncertainty into a set of realizations and for performing sensitivity analysis and optimization. GDSA Framework uses Dakota’s sampling schemes, principally Latin Hypercube Sampling (LHS), to propagate input value uncertainty into probabilistic PFLOTRAN simulations. Dakota is also used in sensitivity analyses to analyze the effects of input value uncertainty on probabilistic GDSA Framework results. In addition, Dakota’s Next Gen Workflow capability was recently incorporated to develop a graphical workflow interface to execute and manage probabilistic GDSA Framework applications (Mariner et al. 2020a).

3. GDSA FRAMEWORK DEVELOPMENT

GDSA Framework has become a powerful PA modeling tool. Over the years, important features and processes for repository PA have been added, several generic repository reference cases have been developed, and probabilistic tools have been established and exercised. These advances were aided by collaboration with other work packages of the SFWST Campaign (and the previous Used Fuel Disposition Campaign) and by interactions with the international community. A recent summary of the major GDSA Framework developments over previous years is provided in Mariner et al. (2019).

Guided by the 5-yr Disposal Research R&D Plan (Sassani et al. 2020), and more distantly the Roadmap reevaluation exercise in FY 2019 (Sevougian et al. 2019), the GDSA team continued to make advances in FY 2021. This section describes advances pertaining to general code development, meshing, uncertainty quantification, workflow, outreach, and international collaboration.

PFLOTRAN advances in FY 2021 are summarized in Section 3.1. Much of this development is described in detail in the recent PFLOTRAN development report (Nole et al. 2021). More general GDSA Framework advances are described in Section 3.2, many of which are also covered in more detail in other reports (e.g., LaForce et al. 2021; Swiler et al. 2021). Section 3.3 addresses additional GDSA infrastructure development and user group activities in FY 2021.

3.1 PFLOTRAN Development

The development of PFLOTRAN for GDSA performance assessment continued at a strong pace in FY 2021. Important advances were made in PFLOTRAN infrastructure (Section 3.1.1), code performance (Section 3.1.2), and process model development (Section 3.1.3). Advances include:

- *Software and hardware infrastructure.* Full adoption and implementation of the Atlassian Jira issue and project tracking software (Section 3.1.1.1); important improvements and additions to the PFLOTRAN QA test suite (Section 3.1.1.2); and increased utilization of new computing resources (Section 3.1.1.3)
- *Code performance improvements.* Characteristic curve smoothing to prevent convergence problems when simulating dry out (Section 3.1.2.1); and implementation of advanced linear and nonlinear solvers to greatly speed up calculations (Section 3.1.2.2). These improvements significantly impact reference cases where phase changes are likely, including the unsaturated alluvium reference case and reference cases subjected to large WP thermal loading.
- *Process model development.* Sharpening of strategies for developing and implementing coupled process models in PFLOTRAN for PA (Section 3.1.3.1); development of an implementation plan for a buffer erosion and waste package corrosion model framework for GDSA (Section 3.1.3.2); continued development of a flexible, PFLOTRAN-compatible fuel matrix degradation (FMD) process model (Section 3.1.3.3; Appendix A); upgrades to the FMD surrogate models of the MATLAB FMD process model (Section 3.1.3.3; Appendix B); enhancements in multi-continuum modeling of coupled fracture-matrix processes (Section 3.1.3.4); implementation of new thermal conductivity models and dependencies (Section 3.1.3.5); implementation and testing of a new smectite-to-illite transition reduced-order model (Section 3.1.3.6); addition of two new criticality-related models, a neutronics surrogate model and a spacer grid degradation model (Section 3.1.3.7); and development of a comprehensive biosphere model (Section 3.1.3.8)

These accomplishments advance GDSA Framework by offering higher fidelity modeling techniques, increasing code stability, and offering a streamlined system for bug reporting and prioritization of future development tasks. Care has been taken to strengthen collaborations across national laboratories and

provide the infrastructure for future collaborations between those laboratories and the international scientific community.

3.1.1 Software and Hardware Infrastructure

3.1.1.1 Code Management System

To help manage code development activities, the PFLOTRAN development team this year fully adopted the Atlassian Jira issue and project tracking software. Jira “issues” range from major upgrades and additions to small tasks and bug corrections. Every two weeks, the team assesses progress made on the issues, categorizes newly submitted issues, re-prioritizes issues as needed, and confirms assignments for the next two weeks.

The extent of adoption of the Jira approach this year is indicated by a large increase in the number of issues in the system. In June of 2020, there were approximately five issues in the system. That number increased to 75 by June 2021.

Figure 3-1 is a snapshot of a Jira dashboard showing summary information for issues in various stages of progress. Details about this system, and how the team utilizes it, are discussed further in Nole et al. (2021).

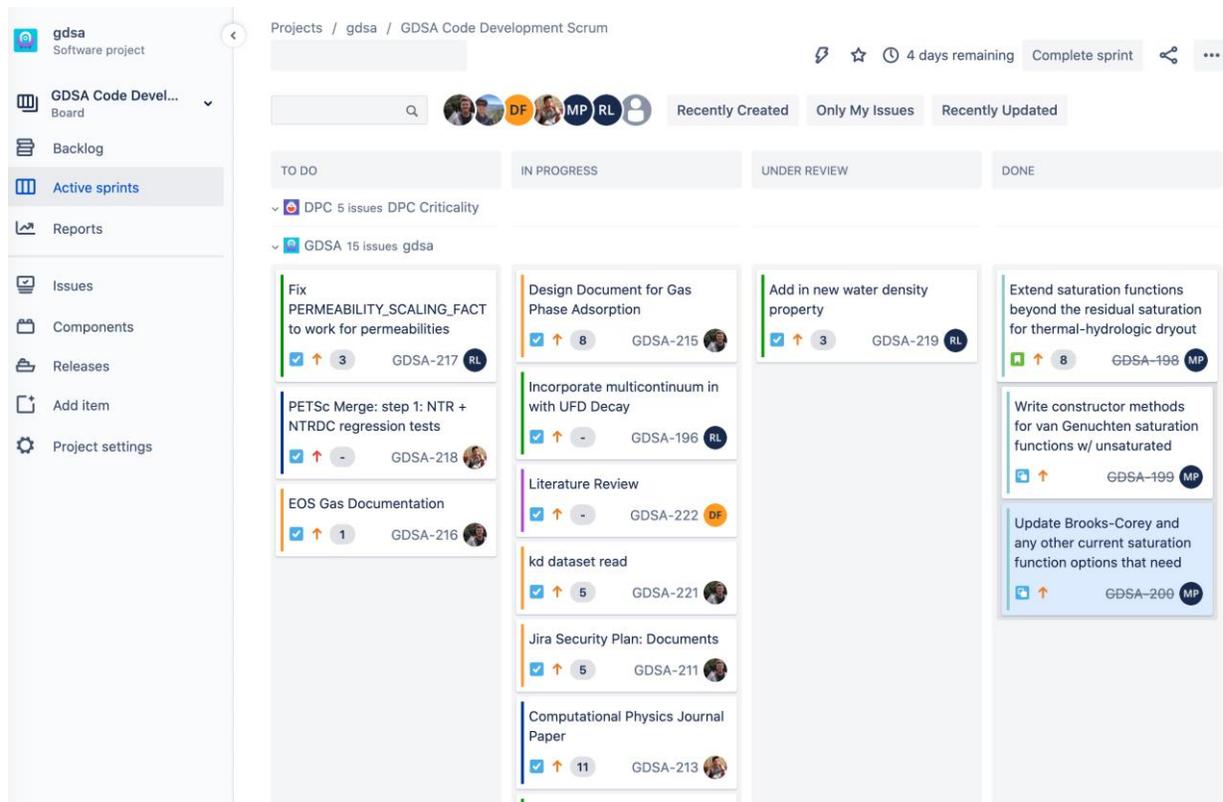


Figure 3-1 Jira summary table snapshot (Nole et al. 2021)

3.1.1.2 Quality Assurance

The GDSA team continued in FY 2021 to expand and upgrade the PFLOTRAN QA test suite. The test suite, established the previous year, was migrated to a new GDSA-QA Git repository where it was fully

exercised. Several of the tests were expanded, and two tests were added, a multi-continuum test and a radial heating problem. Further details of this work are presented in Section 2.1.2 of Nole et al. (2021).

3.1.1.3 *Computing Resources*

Toward the end of FY 2020, the GDSA team acquired a 16-node allocation on a condominium-style 128-node supercomputer called Boca. This allocation allows the GDSA team to run medium- to large-scale simulations without queue time.

The Boca resource has greatly facilitated quick turnarounds for small- and medium-sized problems and quick debugging of input files for the very large probabilistic simulations that are run on the larger supercomputers. More information on this topic is provided in Section 2.1.3 of Nole et al. (2021).

3.1.2 Code Performance Improvements

3.1.2.1 *Characteristic Curve Smoothing*

In FY 2021, characteristic curves for media saturation in PFLOTRAN were smoothed so that PFLOTRAN solvers could handle the transition to dry conditions. Characteristic curves define a porous medium's relationship between water saturation and capillary pressure. Typically, these curves stop or are truncated at the residual water saturation because the curves generally do not consider processes such as evaporation that can cause the water saturation to go below the residual water saturation. Repository simulations, however, must include these conditions because heat from WPs can completely dry out nearby materials.

Prior to this work, transitions to water saturations below the residual were not smooth. For example, a typical transition was a piecewise connection to a maximum capillary pressure. Because Newton solvers require a continuous derivative for all possible water saturations in a simulation, solvers could not make the transition from one piece to the other, and PFLOTRAN simulations could not complete.

Figure 3-2 shows new options implemented this year for the Van Genuchten characteristic curve. Additional details are provided in Section 2.3.1 of Nole et al. (2021). As discussed in LaForce et al. (2021), the smoothing of these curves now allows for full PFLOTRAN simulation of the unsaturated alluvium repository reference case as well as other simulations that involve dry out.

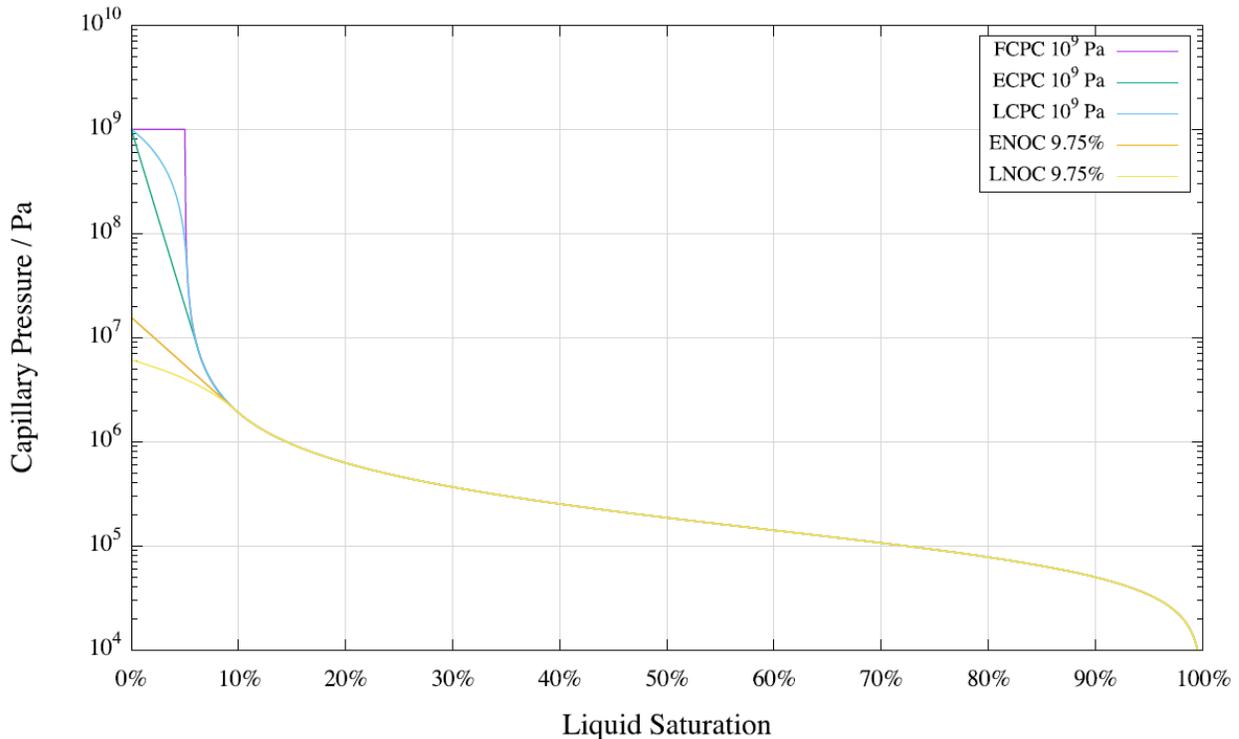


Figure 3-2 Example Van Genuchten characteristic curves employing various extensions at the unsaturated limit (Nole et al. 2021). The purple curve is the original option where the maximum capillary pressure is truncated. The other extension options in the figure are defined using the following abbreviations: F=Flat or constant, E=Exponential, L=Linear, CPC=Capillary-Pressure Cap, NOC=NO Cap.

3.1.2.2 Linear and Nonlinear Solvers

Achieving rapid solver convergence is a continuing challenge in increasingly complicated PFLOTTRAN simulations. Flow and transport calculations rely heavily on both linear and nonlinear solvers and preconditioners. In FY 2021, significant advances were made for both solvers.

For the linear solver, new preconditioners were tested. Preconditioners aim to optimize linear solver performance by generating a more relaxed system of linear equations. Three WIPP nuclear waste repository test cases of increasing complexity (Easy case, Mid case, and Hard case) were used to test new solver preconditioners against the default solver BiCGSTAB (biconjugate gradient stabilized method) for ILU (incomplete lower-upper factorization) preconditioner (BCGS-ILU). The two new solver preconditioners were (1) FGMRES (flexible generalized minimal residual method) for CPR(constrained pressure residual)-Quasi-IMPES(implicit pressure and explicit saturation) (FCQ) and (2) CPR-ABF(alternate block factorization) (FCA). Only the general results are presented here. For details of this analysis, the reader is referred to Section 2.3.4 of Nole et al. (2021).

As shown in Table 3-1 for the Easy case, FCQ was 3 times faster than the default, and FCA was 34 times faster. FCQ and FCA required many fewer time steps and iterations than the default. For the Mid case, FCA had similar results – it was 30 times faster than the default. FCQ was not tested on the Mid and Hard case. For the Hard case, the FCA simulation completed in 41 hours while the default simulation never completed and was ended after 4 days.

Table 3-1 Computation time with Newton-Raphson, Easy case, 32 cores

| | BCGS-ILU | FCQ | FCA |
|----------------------------------|-----------------|------------|------------|
| Computation Time [hours] | 42 | 12.5 | 1.24 |
| Time Steps | 47057 | 10985 | 4633 |
| Nonlinear Iterations | 76874 | 26184 | 15468 |
| Linear Iterations | 28443067 | 5162562 | 227268 |
| Time Step Cuts | | | |
| Linear Solver Fail | 14741 | 2509 | 33 |
| Max Nonlinear Iteration | 284 | 554 | 801 |
| Intentional Time Step Cut | 84 | 294 | 588 |

Regarding nonlinear solvers, the new Newton Trust-Region Dogleg Cauchy (NTRDC) solver, implemented in PFLOTRAN in FY 2020, was also tested on the Mid and Hard cases. On the Mid case when the default linear solver (BCGS-ILU) was used in combination with NTRDC instead of the traditional Newton-Raphson solver, there was a 7-fold increase in speed. The results of the Hard case are shown in Table 3-2. For the Hard case when FCA was used, there was little difference in speed without auto-scaling (AS), but when auto-scaling was used (NTRDC-AS), there was a nearly 3-fold speedup.

Table 3-2 Computation time with FGMRES-ABF, Hard case, 32 cores

| | Newton | NTRDC | NTRDC-AS |
|---------------------------------|---------------|--------------|-----------------|
| Computation Time [hours] | 41.1 | 44.7 | 14.3 |
| Time Steps | 36649 | 37772 | 26433 |
| Nonlinear Iterations | 119362 | 106937 | 118043 |
| Linear Iterations | 17118859 | 16499512 | 3277347 |

Clearly, with large speedups like the ones noted above, pursuing solver optimization is expected to continue to be highly beneficial to GDSA. Details of these tests and others are presented in Section 2.3.4 of Nole et al. (2021).

3.1.3 Process Model Development

3.1.3.1 Strategy for Process Model Enhancements

In FY 2020, a general strategy was developed to facilitate a streamlined quality-controlled process of integrating process models into PFLOTRAN (Mariner et al. 2020a, Section 3.1.3). In FY 2021, that strategy was exercised.

The approach uses the Jira issue and project tracking software discussed in Section 3.1.1.1. Requests for a code enhancement or a new process model are submitted as Jira issues, along with details describing what is needed and the overall importance of the enhancements.

To qualify for the attention of the PFLOTRAN developers, requestors must:

- Provide a summary of the model and general ideas for how it will be implemented
- Identify all model input and output requirements including valid ranges of input values and valid combinations of input values (Detailed requirements on inputs and outputs are listed in Section 3.1.3 of Mariner et al. (2020a).)
- State all model assumptions
- Provide an example simulation of the process model along with plots or tables of outputs

Additional factors considered when prioritizing Jira issues include:

- Sufficient need
 - Do we expect the model to have significant effects on important repository performance metrics or provide important answers to key questions?
- Verification that the process model is ready for implementation
 - Are all model assumptions affecting the validity of the model and the valid ranges of input values acceptable for the intended use?
 - Is there a better approach with more defensible assumptions that covers the same or larger range of applicability?
 - Does the standalone model converge and produce sensible results over the entire multi-dimensional input sample space requested?
- Approval by applicable work package managers

With acceptable answers to the questions above, the requested Jira issue is prioritized and, depending on whether it is high enough priority, assigned to a PFLOTRAN developer. Once assigned, the developer works with the requestor to identify the best implementation approach. Depending on the complexity of the model, the developer may establish regular meetings with the requestor during implementation.

Additional requirements are necessary for the coupling of independent process model codes to PFLOTRAN. The PFLOTRAN repository consists of Fortran 90 simulator source code files and Python libraries for post-processing. It is expected that direct process implementation will be pursued in Fortran and be consistent with PFLOTRAN's style guide. Coupling with external process models requires identifying options for coding languages and any required external libraries so that any potential interface issues can be addressed up front. PFLOTRAN developers will provide recommendations for available options to support database reading/writing. It is expected that process models implemented in PFLOTRAN and codes coupled with PFLOTRAN generate output in PFLOTRAN-consistent formats.

3.1.3.2 Buffer Erosion and Waste Package Degradation

Copper, due to its high chemical stability under crystalline repository conditions, continues to be a prime candidate as a waste package outer barrier material in the U.S. repository program and in the programs of many countries. A concerted effort was made this year to plan PFLOTRAN implementation of buffer erosion and copper corrosion models. The new models are based on the models developed and used by SKB for the Forsmark repository and used by Posiva for Olkiluoto. Adding this type of modeling capability to PFLOTRAN is expected to be useful not only for evaluating buffer erosion and the performance of copper in the U.S. program but also potentially to facilitate adoption and testing of PFLOTRAN for repository applications outside the U.S.

Conceptual Model

The chronology of the proposed conceptual buffer erosion copper corrosion model is:

- Buffer in a deposition borehole (or tunnel) erodes due to groundwater flow from an intersecting fracture, as depicted in Figure 3-3
- Buffer erosion eventually exposes the copper surface of the WP to flowing groundwater
- Hydrogen sulfide in the flowing groundwater attacks the exposed copper

In effect, the conceptual model involves two serially coupled models, buffer erosion followed by copper corrosion.

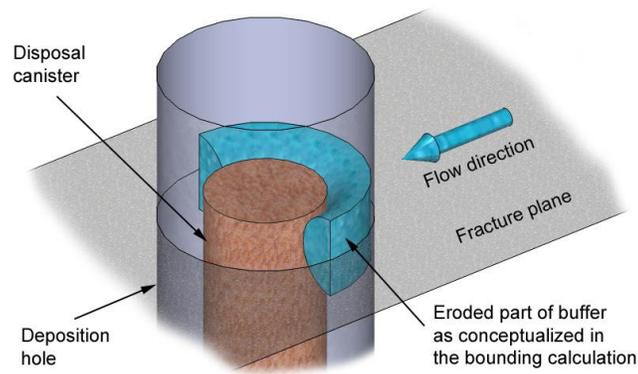


Figure 3-3 Conceptual model of buffer erosion due to a flowing fracture (Fig 6-108, Posiva 2013)

Details of the conceptual model for buffer erosion are based on the chemical erosion model summarized in Posiva (2013, Section 6.11). Assumptions include:

- Clay particles are removed from the borehole or tunnel due to swelling-induced extrusion into fractures and the shear forces of flowing groundwater or due to dispersion as a colloidal sol
- Buffer consists primarily of Na-montmorillonite
- Dilute water (2-4% total dissolved solids, or ionic strength below 10 mM) is needed to stabilize Na-montmorillonite colloid dispersions, allowing buffer erosion

Assumptions regarding repository design include:

- The repository host rock is fractured crystalline rock
- Bentonite buffer surrounds canisters in deposition holes and fills the voids in the holes (or surrounds canisters emplaced in tunnels and fills the tunnel voids)
- Optional: DRZ surrounds the deposition holes (emplacement tunnels)

Mathematical Model

The shape of buffer erosion in the deposition hole (or tunnel) is assumed to be a quarter donut, i.e., one side of the hole is eroded as indicated in Figure 3-3. The model calculates the amount of additional buffer eroded during each time step and adds it to the cumulative amount eroded.

The buffer erosion rate is determined using the relationship shown in Figure 3-4, which is described in Section 6.11 of Posiva (2013). This rate is an empirical relationship that is a function of flow velocity of the intersecting fracture, fracture aperture, and the emplacement geometry.

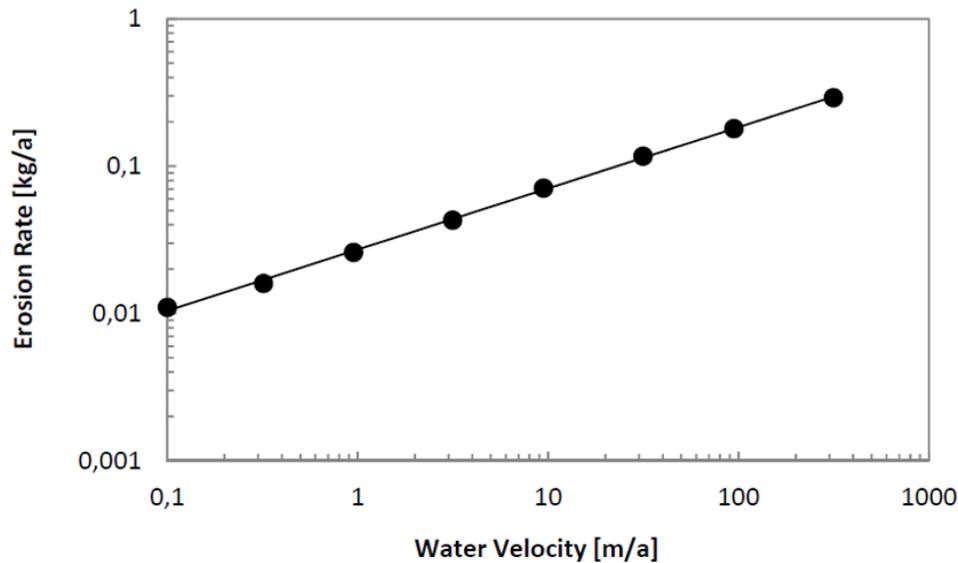


Figure 3-4 Montmorillonite release rates as a function of water velocity in the intersecting fracture (Posiva 2013, Figure 6-97)

Once the cumulative volume of the eroded buffer exceeds the volume of the quarter donut, the buffer is effectively breached and the copper corrosion model activates. The surface area of the exposed waste package is assumed to be the area of a half band around the canister having a width equal to the distance between the borehole wall and canister (Posiva 2013, Eq. 6-112).

The relationship in Figure 3-4 is described by the following equation:

$$r = a\delta v^b \quad (3-1)$$

In this equation, r is the buffer erosion rate (mass per time), δ is the fracture aperture, and v is the water velocity in the fracture. The constants a and b depend on the application. For example, values of 27.2 and 0.41, respectively, can be used for a 1 mm fracture aperture, 1.75 m deposition hole diameter, water velocities from 0.1 to 315 m yr⁻¹, and a 0.4 montmorillonite volume fraction at the fracture/deposition hole interface (Posiva 2013). For erosion of backfill in a 5-m diameter tunnel values of 20.7 and 0.598, respectively, can be used for these constants (Eq. 6-76). It is up to the users to determine the proper values for specific applications.

The copper corrosion model is represented mathematically just as it is in the SKB model (Posiva 2013, Section 6.19). An equivalent groundwater flow rate (volume per time) from the intersecting fracture is calculated from the lesser of two relationships, one of which accounts for diffusion. For every two moles of hydrogen sulfide flowing into the exposed region, one mole of copper corrodes and precipitates as CuS. As soon as the exposed copper fully corrodes, the WP is breached.

Implementation Plan

PFLOTRAN implementation of this model needs to accommodate parallel processing, stochastically-generated DFNs, and an equivalent continuous porous medium (ECPM) mesh. To be successful, it needs to:

- Automate the identification of fractures intersecting the deposition holes (or emplacement tunnels)
- Automate the acquisition of flow velocities, fracture apertures, and concentrations of Na, Ca, and hydrogen sulfide from the identified intersecting fractures
- Adjust the properties of buffer cells affected by erosion to be consistent with the SKB buffer erosion model

In general, python scripts will be used to identify and characterize fractures intersecting deposition holes, “faux rock” simulations will be used to determine flow velocities and concentrations of Na, Ca, and hydrogen sulfide in intersecting fractures over time, and a separate buffer cell erosion model will be developed to emulate the buffer erosion predicted by the SKB buffer erosion rates. These plans are discussed in more detail below.

Because parallel processing complicates spatial coupling, flow velocities and water compositions in nearby fractures cannot be readily used to calculate buffer erosion and copper corrosion rates for a given waste package and deposition hole. To avoid this problem, initial “faux rock” simulations will be performed in which each deposition hole is returned to its original (pre-drilled) fracture rock state so that fracture flow velocities and water compositions can be captured through time at each cell where a fracture would exist in the deposition hole domain. This approach allows for simple parallel processing by eliminating the need for spatial coupling to parameter values in other cells. This approach is expected to provide a reasonable approximation of the fracture flow velocities and groundwater concentrations that would occur in the presence of the deposition holes. To help make these fracture velocities and water compositions as accurate as possible, the tunnels and their backfill and the tunnel DRZ will not be replaced with faux rock, and a heat source will be included where the waste packages would be. In addition, the thermal conductivity of the faux rock cells may be adjusted to mimic that of the waste package and buffer.

The SKB buffer erosion model is a zero-dimensional model, whereas the repository reference case model domain is three-dimensional. The model domain has cells containing buffer that should lose buffer as the zero-dimensional model predicts. Consequently, two distinct buffer erosion models are planned for implementation, a *virtual buffer erosion model* (the SKB model) and a *buffer cell erosion model*. The SKB model is referred to as the virtual buffer erosion model because it will be run in the waste package domain and will not itself remove any material from the model domain. The *buffer cell erosion model* will be run at each buffer cell location where there is a qualifying intersecting fracture. It will be made to be consistent with the virtual buffer erosion model and will remove buffer from the domain, thereby enhancing the flow of water through the deposition hole and the flow of corrosive reactants to the waste package surface.

The erosion models and the groundwater flow and transport simulations are not directly coupled to avoid complications in parallel processing. The “faux rock” simulation will run the entire simulation time and is required to output the following information to run the *virtual buffer erosion model* or the *buffer cell erosion model*:

- The groundwater flow velocity, Na concentration, and Ca concentration history per time step at each WP location.

- The identification number and coordinates of each WP corresponding to the time history data.

The second “full” simulation will require these data as inputs to simulate the buffer erosion and copper corrosion models.

There are several required inputs to run the buffer erosion models:

- Fracture flow velocity (m yr^{-1}), Na concentration, and Ca concentration at each coordinate (or waste package ID) per simulation time (that is saved at each time step) will either be provided by the output file from the “faux rock” simulation or manually by the user in the same format. The manual option will be useful when testing the model.
- The fracture apertures at each waste package borehole location will be extracted from DFN realizations or inferred from ECPM realizations. This information should be consistent with the “faux rock” simulation for accurate buffer erosion results. This information can also be generated manually by a user. The location of waste package and buffer will be defined in the input of PFLOTRAN for both the “faux rock” and the full model simulations.
- The buffer erosion rate equation parameters are inputs in the PFLOTRAN input deck as a dataset. These values can be individually calculated per location depending on the fracture aperture from DFN or ECPM. The other option is to enter these values per location manually by the user.
- When Na and Ca concentrations cross a user-provided threshold, the buffer erosion process either halts (above threshold) or commences (below). Na and Ca concentration thresholds are defined for the simulation. Suggested values from SKB will be defined as default values; otherwise, a user can enter it in the PFLOTRAN input deck. The borehole radius, canister radius, buffer grain density and porosity, and buffer porosity will be defined in the MATERIAL_PROPERTY and GRID of the PFLOTRAN input deck.

The virtual buffer erosion model will be its own module or a part of the waste package module in PFLOTRAN. The capability to remove buffer from buffer cells in accordance with the buffer cell erosion model is already built into PFLOTRAN. The copper corrosion model will turn on for the waste package when the total buffer eroded exceeds the volume of the quarter donut.

An implementation request for the buffer erosion and copper corrosion models has yet to be submitted to the Jira system. In accordance with the implementation guidelines in Section 3.1.3.1, more work is needed to ensure and document the need and feasibility of the proposed approach. It is expected, however, that by early FY 2022 the implementation request will be submitted and the decision on whether to proceed will be made.

3.1.3.3 Fuel Matrix Degradation (FMD)

The FMD model is the uranium dioxide (UO_2) matrix degradation process model of GDSA Framework. It was developed collaboratively at Argonne National Laboratory and Pacific Northwest National Laboratory (Jerden et al. 2015). The model, coded in MATLAB, calculates spent fuel degradation rates as a function of radiolysis, electro-kinetic reactions, alteration layer growth, and diffusion of reactants through the alteration layer. In more recent versions of the model, steel corrosion is included to provide a source of hydrogen (Jerden et al. 2018).

In FY 2015 a version of the FMD process model that excludes steel corrosion was coded in Fortran (Jerden et al. 2015). That Fortran code was coupled to PFLOTRAN and successfully demonstrated (Mariner et al. 2015). Apart from the computational results, the demonstration indicated that the coupled model was computationally demanding. Mechanistic simulation of the FMD model processes requires many calculations at each time step. For a probabilistic repository PA calculation there are thousands of

WPs and hundreds of realizations. Though the coupled Fortran FMD process model may be used for small scale problems, it is too expensive for higher-fidelity repository PA simulations that can propagate spatial variability in the values of inputs (e.g., local temperature and local environmental concentrations of chemical reactants) across the repository.

In FY 2019 and FY 2020, two approaches were undertaken to include the FMD process model in GDSA Framework. One approach, summarized in the first subsection below, was to develop surrogate models of the FMD process model for accurate and rapid emulation in PFLOTRAN. The other approach, described in the subsequent subsection, was to develop a new Fortran code that utilizes rapid solvers and flexible time steps for a more powerful standalone FMD process model that will smoothly couple to PFLOTRAN.

FMD Surrogate Models

In FY 2020, an artificial neural network (ANN) surrogate model and a k-Nearest-Neighbors regression (kNNr) surrogate model were implemented in PFLOTRAN for the FMD process model. The ANN surrogate is a parametric model that utilizes a network of artificial neurons with nonlinear activation functions. The kNNr surrogate is a nonparametric model that uses an advanced technique to approximate the response from the closest set of neighboring training points in a multidimensional database. Both approaches were pursued to increase the chances of developing a useful surrogate and to allow comparison of the advantages and disadvantages of the approaches for incorporating surrogate models in GDSA Framework. That work was documented in Mariner et al. (2020).

In FY 2021, a modified sampling strategy was developed and used to generate training data for the ANN and kNNr surrogate models. This work is documented in detail in Appendix B. There were approximately eight times as many simulations in the parameter space. In addition, down sampling by a factor of ten in the time dimension led to a more balanced data density across all dimensions in the sample space. For kNNr, this new data set together with a narrowing of the input features from six to four led to an improvement of the kNNr mean square error (MSE) over the test set by a factor of 2 compared to Mariner et al. (2020b). Besides this factor of two improvement, the updated approach indicates that adding yet even more samples in the configuration space will continue to improve the accuracy of kNNr.

Table 3-3 compares the accuracy of ANN and kNNr surrogates on the test data set. For the new data sets, the kNNr method is slightly more accurate in the MSE metric than the ANN method. The ANN surrogate shows better accuracy in the mean absolute percentage error (MAPE). In the mean absolute error (MAE) metric, the ANN is slightly more accurate than kNNr. Thanks to the improved sampling approach, the MSE error in the optimal kNNr approach is less than half of what it was in Mariner et al. (2020b).

Table 3-3 Values of the error metrics computed on the test set (and training set for the ANN) for the UO₂ surface flux in the “natural” (i.e., non-log₁₀ transformed) space. The error on the training data is zero for kNNr because for inverse-distance weighted averaging, the table prediction is the same as the tabulated value at the query point.

| Surrogate | Train MSE (mol m ⁻² yr ⁻¹) ² | Test MSE (mol m ⁻² yr ⁻¹) ² | Train MAE (mol m ⁻² yr ⁻¹) | Test MAE (mol m ⁻² yr ⁻¹) | Train MAPE | Test MAPE |
|--------------------|---|--|--|---|------------|-----------|
| ANN | 3.41e-6 | 3.08e-6 | 8.57e-4 | 8.26e-4 | 33.7% | 25.1% |
| kNNr 2.28M/80NN | N/A | 2.69e-6 | N/A | 8.92e-4 | N/A | 44.4% |

As in Mariner et al. (2020b), the surrogate results were found to be highly accurate where training data were densely spaced. This finding underscores the importance of focusing surrogate model training data on the expected sample space of the application. Further, surrogate simulation was found to be rapid compared to the coupled process model, allowing full repository simulation of the shale repository reference case in a practical amount of time. These results indicate that each surrogate model will enable GDSA Framework to rapidly and accurately simulate spent fuel dissolution for each individual breached spent fuel WP in a probabilistic repository simulation.

Figure 3-5 demonstrates the new FY 2021 ANN FMD surrogate results in a shale reference case simulation with emplacement of both 24 pressurized water reactor (PWR) and 37 PWR WPs. The chemical conditions for this simulation are: 1×10^{-9} M O_2 (aq), 1×10^{-5} M H_2 (aq), 3×10^{-6} M Fe^{2+} , 1.4×10^{-5} M CO_3^{2-} . For comparison, the figure includes the results of the traditional fractional dissolution rate (FDR) model that uses a dissolution rate of 10^{-7} yr⁻¹. Waste packages start to breach in these simulations around 300 years. Note that the surface-area-specific degradation rate (bottom plots) decreases over time in the ANN FMD simulation, as would be expected due to decreases in dose rate and temperature with time. For the FDR model, a surface-area-specific rate of $\sim 3.7 \times 10^{-7}$ mol m⁻² yr⁻¹ is maintained after breach for the rest of the simulation (both simulations assume a constant fuel specific surface area of 0.001 m² g⁻¹).

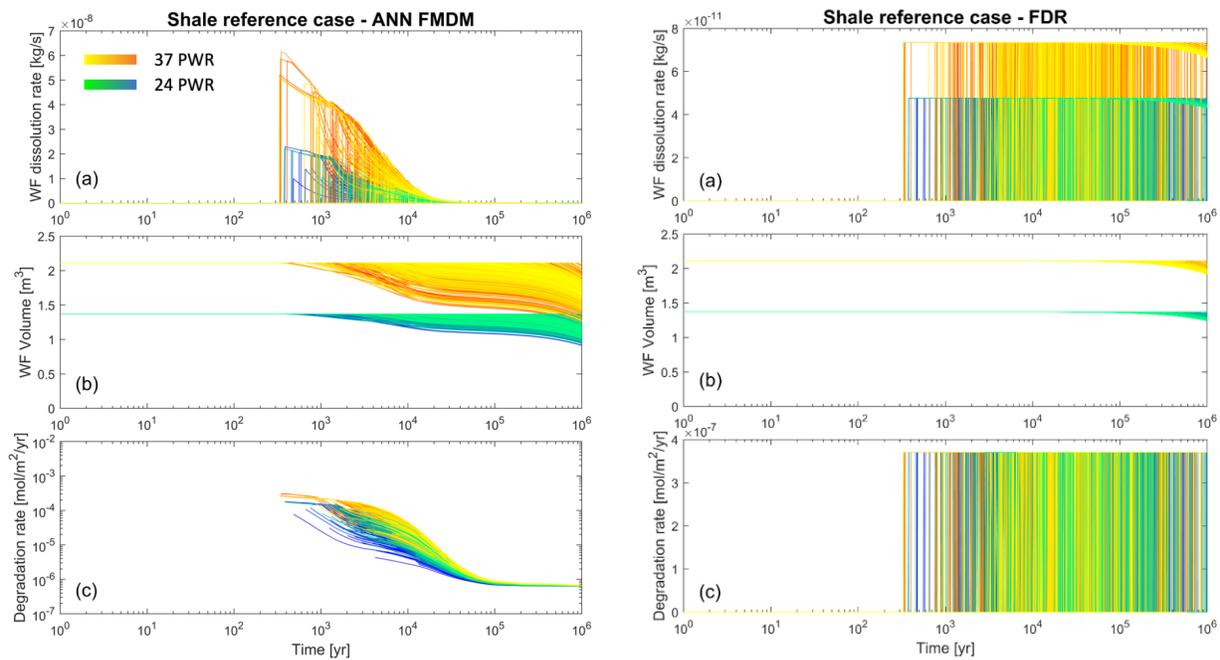


Figure 3-5 Fuel matrix degradation rates (a) and (c) and remaining fuel volumes (b) for breached 24 PWR (green and blue) and 37 PWR (yellow and orange) waste packages in a shale repository reference case using the ANN FMD surrogate (left) and traditional fractional dissolution rate (10^{-7} yr⁻¹) model (right).

Fortran Process Model

The original Fortran FMD code developed in FY 2015 and coupled with PFLOTRAN uses much of the same structure as the MATLAB code. It recursively solves for corrosion potentials and new concentrations instead of coupling the solvers which leads to an increase in iterations and decrease in speed. Further, the 2015 Fortran code was not improved to allow for dynamic time stepping and was not set up for parallel computing.

Under the continued guidance of Glenn Hammond, development of a new Fortran code is underway for faster run times, smoother PFLOTRAN coupling, and parallel computing. The goals of this initiative are to develop a faster FMD process model with flexible coupling features, to keep it open source, and to help the process modelers transition to the new code so that they can continue to develop a code that is compatible with PFLOTRAN.

Appendix A documents progress in FY 2021 toward developing an improved solution scheme. Initial tests indicate that the number of iterations required to converge to a solution will be significantly reduced. This will likely provide a non-trivial speed-up, potentially reduce the likelihood of discontinuities in the observed fluxes and help avoid the non-convergence issues often encountered with the MATLAB code.

Efforts this year enabled implementation of a dynamic time stepping algorithm, studying control parameters within the FMD model, and understanding the role of 5th order reactions incorporated into the model. Long time periods are frequently needed to reach solution equilibrium, therefore the dynamic time stepping routine is needed to reach simulation times of 1000 years as indicated in Figure 3-6. Moreover, the MATLAB implementation works on a fixed logarithmic time step which could potentially lead to discontinuities if concentrations fluctuate on short time scales. Our algorithm allows for the time step to be shortened, as well as increased, which will help resolve these challenges. An exploration of the control parameters and 5th order reaction rate expressions within the model indicate that reactions can be rapidly turned off as the temperature and concentrations within the rate expressions decrease. This can cause the model to produce non-intuitive results that are difficult to predict. Smoother rate expressions could be implemented in the future, although we emphasize that the model has compared well with experiments. On-going efforts are focused on interfacial reactions which provide the oxidative dissolution of the fuel surface and are the source of the excessive iterations mentioned above.

Appendix A outlines in detail the conceptual model, developing code, and recent progress.

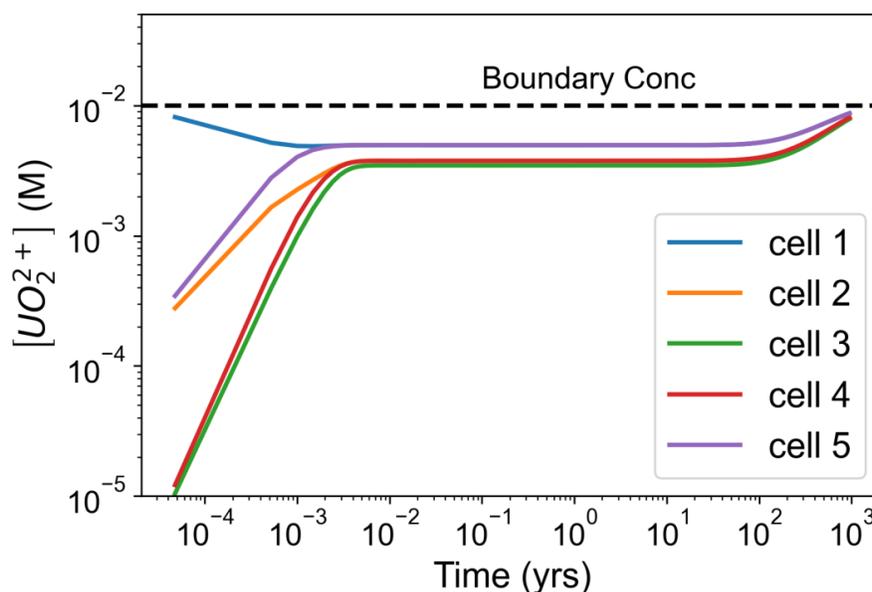


Figure 3-6 Concentration (M) of UO_2^{2+} as a function of time in a log-log plot. A long equilibration time (~100 yrs) and > 1000 yrs is needed to reach the boundary condition concentration. A dynamic time stepping routine was implemented to allow for long total simulation times.

3.1.3.4 Multi-Continuum Transport

PFLOTRAN's multi-continuum model simulates disconnected secondary (matrix) continua connected to a primary (fracture) continuum. It is referred to as the DCDM (Dual Continuum Disconnected Matrix) model (Lichtner, 2000). Each primary cell hosts its own disconnected one-dimensional secondary continuum.

In recent years, the multi-continuum model has been enhanced in several ways. Sorption capabilities, dispersion in the fracture, spatial aperture discretization, and parallel capabilities have all been improved. In FY 2021, this model was verified against two analytical solutions and applied to a three-dimensional discrete fracture network (DFN) problem (Nole et al. 2021, Sections 2.2.1.6, 2.2.1.7, and 2.2.1.8).

One of the analytical solutions tested in FY 2021 is from Tang et al. (1981). It describes advective and diffusive transport in a single fracture with diffusion into a semi-infinite rock matrix. Transport was successfully predicted in all cases. An example comparison of PFLOTRAN simulations versus the analytical solution is shown in Figure 3-7.

The other analytical solution tested was that of Sudicky et al. (1982). This solution is an extension of the Tang et al. (1981) solution. The difference is that this one pertains to a rock matrix of finite length. Two matrix lengths were tested, 0.05 m and 0.25 m. As in the case of the Tang et al. (1981) solution, a high degree of accuracy was observed. Details of these analyses are presented in Section 2.2.1 of Nole et al. (2021). These results indicate that the simplifying assumptions of the Dual Continuum Disconnected Matrix multi-continuum model do not introduce much model error.

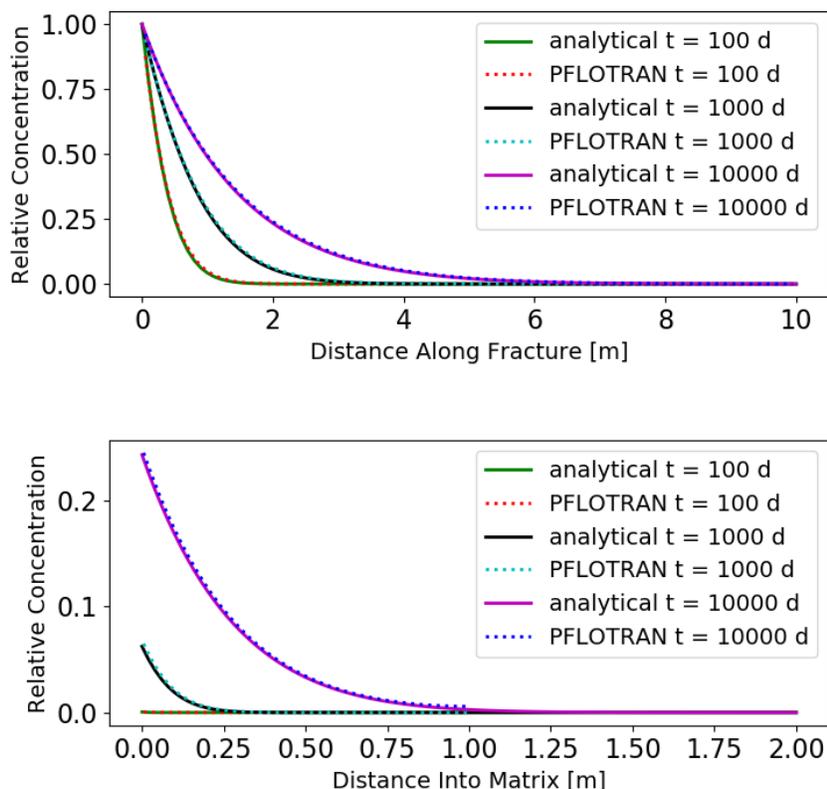


Figure 3-7 Comparison of PFLOTRAN results to the analytical solution of Tang et al. (1981) along the fracture (top) and into the matrix (bottom) at 2 m along fracture

3.1.3.5 Thermal Conductivity Enhancements

Accurate simulation of thermal conductivity is needed for accurate prediction of heat dissipation, especially for criticality analysis. Again this year, there were major advances in the modeling of thermal conductivity in PFLOTRAN. The advances began last year with the development and implementation of thermal characteristic curves (TCCs) for specific materials. That work continued in FY 2021 along with the implementation of anisotropy and other special models.

Thermal Characteristic Curves

Prior to FY 2020, there was only one TCC implemented in PFLOTRAN. That “default” TCC is from Somerton et al. (1974) and is a function of wet and dry thermal conductivity and liquid saturation. In FY 2020, four new thermal conductivity characteristic curves were implemented: constant, linear resistivity, cubic polynomial, and power law (Mariner et al. 2020a, Section 3.1.9).

In FY 2021, a permafrost TCC was added. Modeling the thermal conductivity for permafrost is expected to be useful for modeling heat dissipation during glacial periods. Also, in FY 2021, the overall TCC feature was upgraded in PFLOTRAN to provide greater flexibility and improved error messaging. Details of these advances are provided in Section 2.2.3 of Nole et al. (2021).

Thermal Conductivity Anisotropy

Thermal conductivity in a material is often anisotropic. For example, layered sediments typically exhibit conductive (thermal or hydraulic) anisotropy in the layered directions versus the transverse direction. Engineered features, including the assemblies within a WP, may also exhibit anisotropic characteristics.

Thermal conductivity is implemented as a symmetric tensor with six unique components and unit vectors. The off-diagonal elements are zero when the grid aligns with the principle axes of conductivity. The new capability includes error messaging for possible mistakes the user could make when using this feature. In addition, the capability of generating a composite anisotropic curve from curves defined for the principle directions is under development. Details are provided in Section 2.2.4 of Nole et al. (2021).

Special Thermal Conductivity Models

Specific thermal models for DPCs have been developed for implementation in PFLOTRAN. There are different models for axial and radial directions and wet and dry conditions. The axial and radial models may be combined into a composite anisotropic model. These models are described in Section 2.2.5 of Nole et al. (2021).

3.1.3.6 Smectite-to-Illite Transition

In FY 2021, a new smectite-to-illite (illitization) reduced order model was developed and implemented in PFLOTRAN. This model can simulate the gradual change in the ratio of montmorillonite to illite in a bentonite buffer and the corresponding expected change in permeability.

The rate of change in the smectite fraction is based on a relationship observed by Huang et al. (1993):

$$-\frac{df_S}{dt}\Big|^{i+1} = \begin{cases} [K^+] \cdot (f_S^i)^2 \cdot A \exp\left(-\frac{E_a}{\mathcal{R} \cdot T^{i+1}}\right) & T^{i+1} \geq T_{th} \\ 0 & T^{i+1} < T_{th} \end{cases} \quad (3-2)$$

For a given time step $i+1$, the rate of change is a function of the potassium cation concentration $[K^+]$ in mol L^{-1} and the previous smectite fraction f_S^i . A is the frequency term in L/mol-s , E_a is the activation energy in J mol^{-1} , \mathcal{R} is the ideal gas constant, T^{i+1} is the temperature in Kelvin of the grid cell, and T_{th} is the threshold temperature below which the reaction does not take place (Nole et al. 2021, Section 2.2.6).

The new permeability k_j^{i+1} is computed using the following relationship:

$$k_j^{i+1} = k_j^0 \left[1 + \left(\frac{f_I^{i+1} - f_I^0}{f_S^0} \right) \cdot C_k \right] \quad (3-3)$$

where k_j^0 is the original permeability tensor, C_k is a shift factor, $(f_I^{i+1} - f_I^0)$ is the change in the illite fraction, and f_S^0 is the original smectite fraction. Further details are provided in Nole et al. 2021, Section 2.2.6).

This model was tested on a simple three-dimensional domain with two different buffer regions as shown in Figure 3-8. During the simulation, smectite nearly completely transitioned to illite and the permeability adjusted accordingly. Time histories of temperature, mineral fraction, and permeability for buffer region #2 are shown in Figure 3-9. Details of this test problem are included in Nole et al. (2021, Section 2.2.6).

Future work will involve expanding the code into a more flexible generalized PFLOTRAN module and adding additional illitization models (Nole et al. 2021, Section 2.2.6).

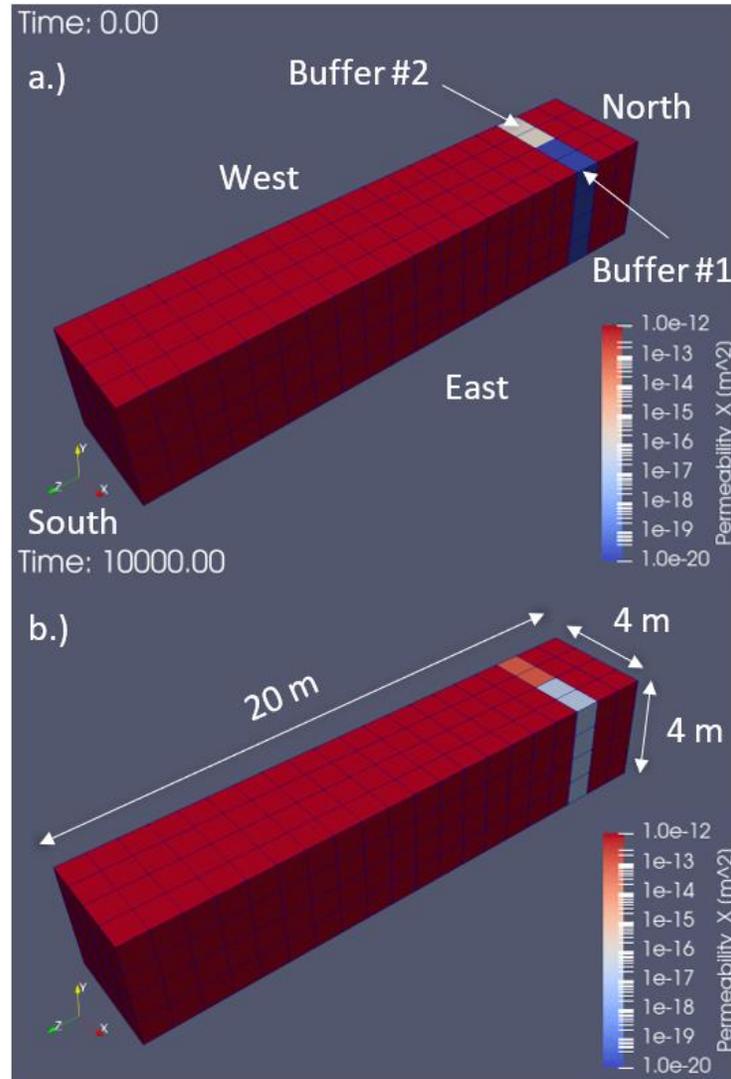


Figure 3-8 Smectite to illite evolution at 0 years (a) and 10,000 years (b) of buffer placed 2 to 3 meters from the north end of the domain. Buffer 1 is originally 90% smectite and has a reaction threshold temperature of 40°C. Buffer 2 is originally 75% smectite and has a reaction threshold temperature of 60°C. The rest of the domain is water-saturated soil with a porosity of 25%, tortuosity of 50%, and permeability of 10^{-12} m². Initially the domain has a hydrostatic pressure of 1 MPa and a temperature of 20 °C. At the beginning of the simulation, the north face boundary condition is set to 5 MPa and 260 °C.

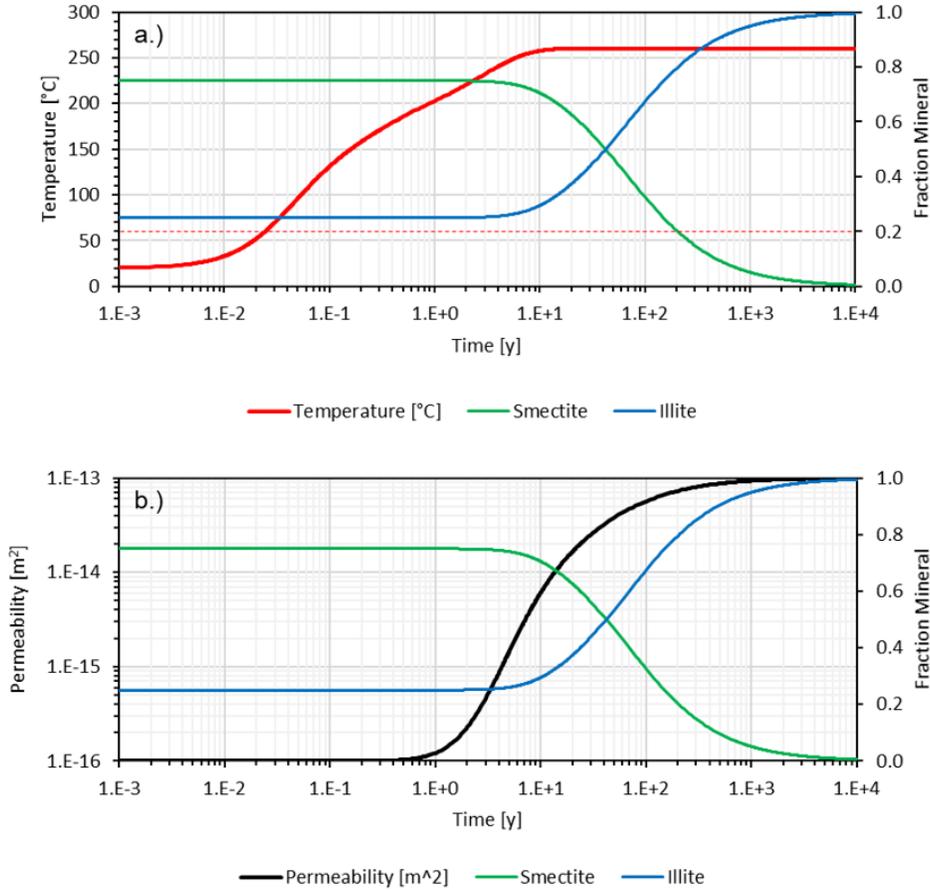


Figure 3-9 Smectite to illite time histories showing thermal and permeability changes for Buffer #2

3.1.3.7 Criticality-Related Models

Two new criticality-related models were added to PFLOTRAN in FY 2021, a neutronics surrogate model and a spacer grid degradation model. Only brief descriptions of these models are provided here. Section 2.2.7 of Nole et al. (2021) describes these models in detail and presents two test applications.

The neutronics surrogate model uses lookup tables to provide decay heat, heats of criticality, and evolving radionuclide inventory (if chosen). The values in the lookup tables are generated a priori from coupled neutronics and thermal-hydraulic simulations applied over expected ranges of hydrological conditions (e.g., temperature, water saturation, water density). This surrogate model is currently applicable to steady-state criticality events.

The spacer grid degradation model tracks the corrosion of the spacer grid over time. When the grid is sufficiently corroded, the grid loses its vitality and the fuel rods lose their critical configuration, terminating the possibility for criticality. The corrosion rate for the spacer grid is a function of a corrosion rate constant \mathcal{C} , activation energy Q , average WP temperature \bar{T} , cumulative spacer grid surface area A_0^{SG} , and a saturation-dependent ratio $f(\bar{S}_i)$:

$$R^{i+1} = f(\bar{S}_i^{i+1}) \cdot A_0^{SG} \cdot \mathcal{C} \exp\left(-\frac{Q}{\mathcal{R} \cdot \bar{T}^{i+1}}\right) \quad (3-4)$$

These two models were designed to work in concert with each other and with the smectite-to-illite transition model (Section 3.1.3.6).

The neutronics surrogate and spacer grid degradation models were executed together in the test case shown in Figure 3-10. Three waste forms (WFs) were emplaced with unique breach times, WF degradation mechanisms, and criticality mechanisms. WFs #1 and #2 invoked the spacer grid degradation model. Critical water saturations were specified for each WF, and critical water densities were specified for #1 and #3. Details of this test case are provided in Section 2.2.7.2 of Nole et al. (2021).

Results are shown in Figure 3-11. The spacer grid vitality was lost at 43 and 36 days for WFs #1 and #2, respectively, not long after the start of the criticality periods. Consistent with the start of the criticalities, the temperatures spiked. These results and others are discussed in detail in Section 2.2.7.2 of Nole et al. (2021).

These new criticality-related models are a major advance in our ability to directly include criticality processes and events in total system performance assessment calculations. With the ability to simulate criticality in repository simulations, we can better predict whether criticality can occur and better understand the coupled effects of heat, water saturation, EBS material performance, radionuclide release, and radionuclide transport in the near field.

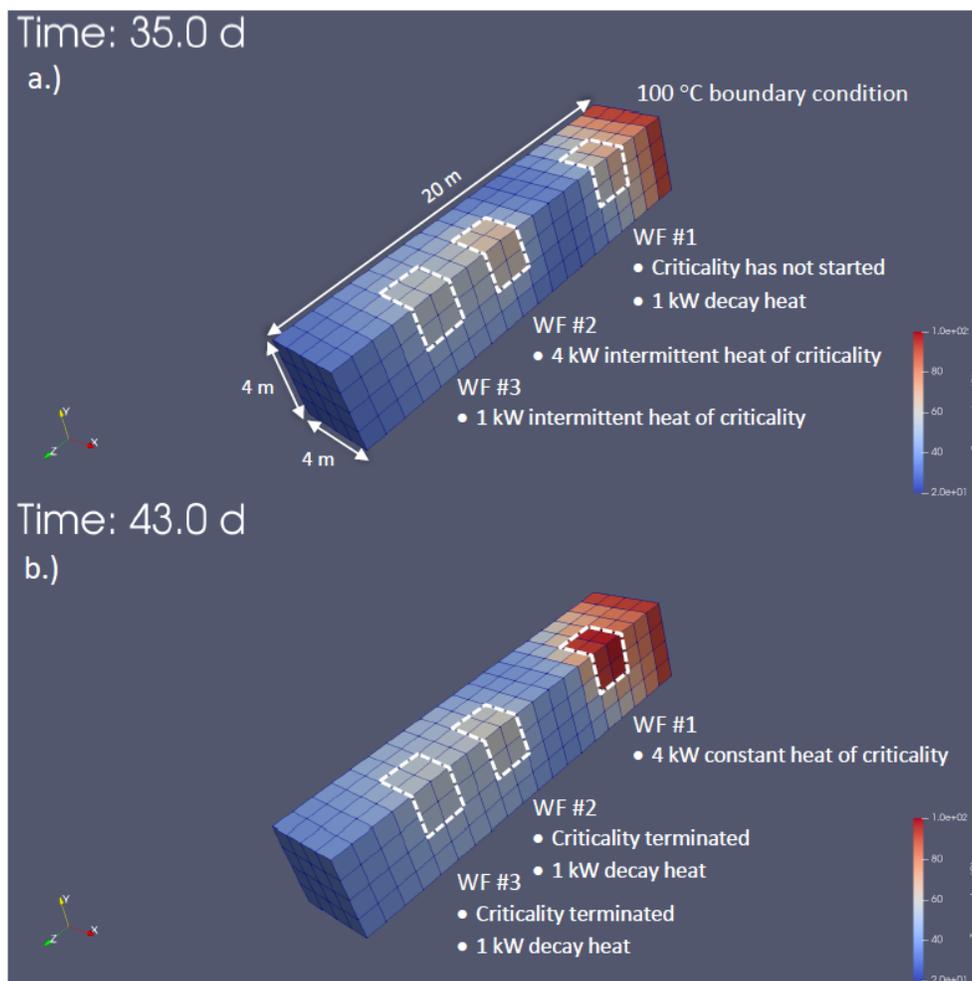


Figure 3-10 Neutronics surrogate and spacer grid degradation test problem showing temperatures at 35 days (a) and 43 days (b) for a domain with three different waste forms with unique criticality events. At 35 days WF #1 has not reached criticality while WF #2 and #3 have. At 45 days criticality has ceased at WF #2 and #3 and begun at WF #1.

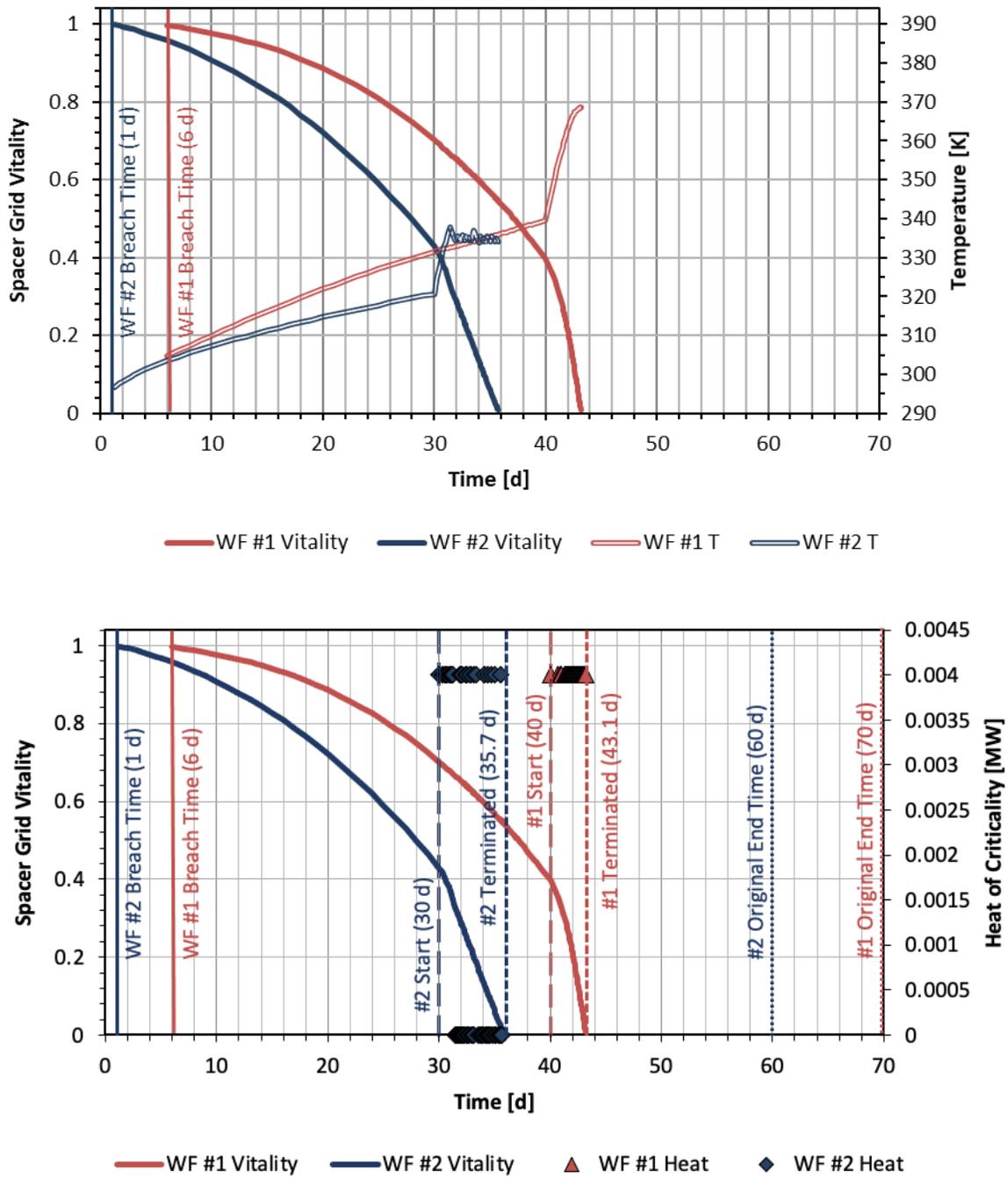


Figure 3-11 Spacer grid vitality, temperature, and heat of criticality over time showing loss of spacer grid vitality at 43 and 36 days for WFs #1 and #2, respectively.

3.1.3.8 Biosphere

The current biosphere model implemented in PFLOTRAN consists of an ingestion dose model from drinking contaminated well water (Mariner et al. 2017, Section 3.2.3). That model can be used to simulate Example Reference Biospheres 1A and 1B of the International Atomic Energy Agency (IAEA 2003) and can explicitly include the effects of highly-mobile short-lived radionuclides like radon-222.

In FY 2020, a GDSA work package at Pacific Northwest National Laboratory (PNNL) was created to plan the development of a comprehensive biosphere model for GDSA Framework. The general requirements of the new model are that it be generic, flexible, open source, compatible with PFLOTRAN, and consistent with international recommendations and guidance for such models built for deep geological repositories. The new biosphere model will account for multiple pathways, features, and interactions and will include radionuclide decay and ingrowth.

Figure 3-12 shows a schematic diagram of the various pathways planned for inclusion. Figure 3-13 shows the general interaction matrix. A detailed description of the plans for this model is provided in Condon et al. (2020). A full report on the progress made in FY 2021 will be available in September of 2021.

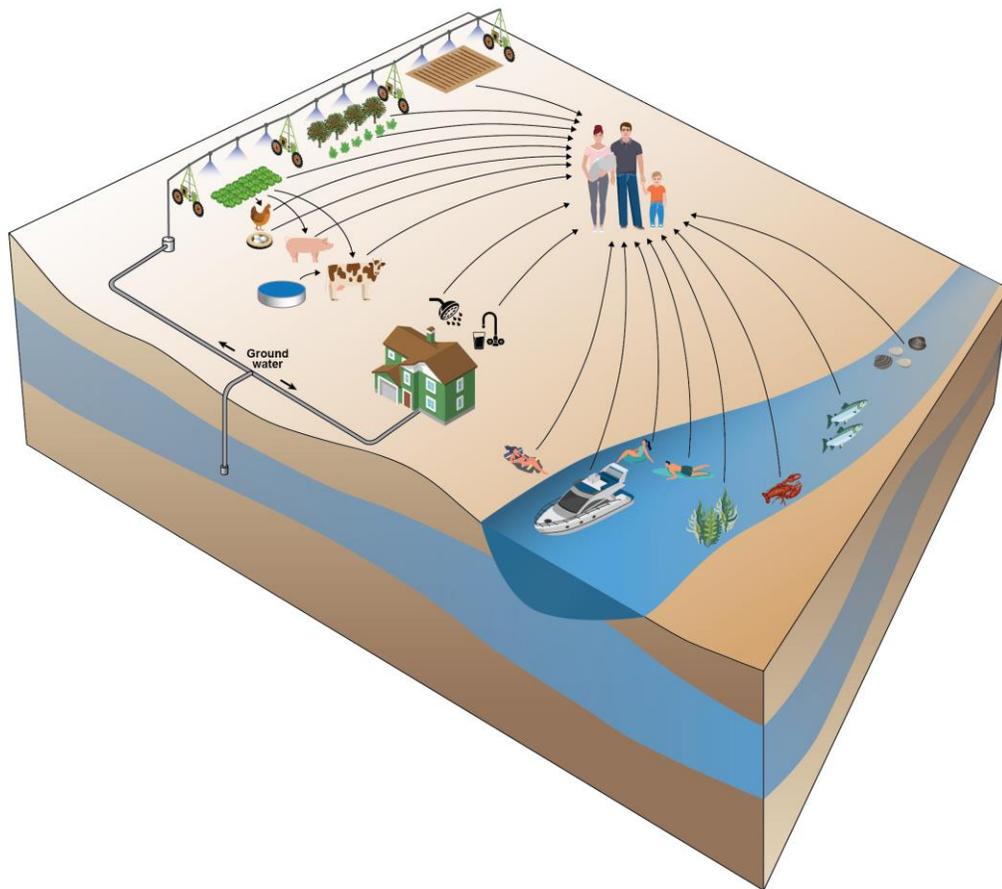


Figure 3-12 Schematic diagram of pathways to be included in new biosphere model being developed for GDSA Framework (graphic from Condon et al. 2020)

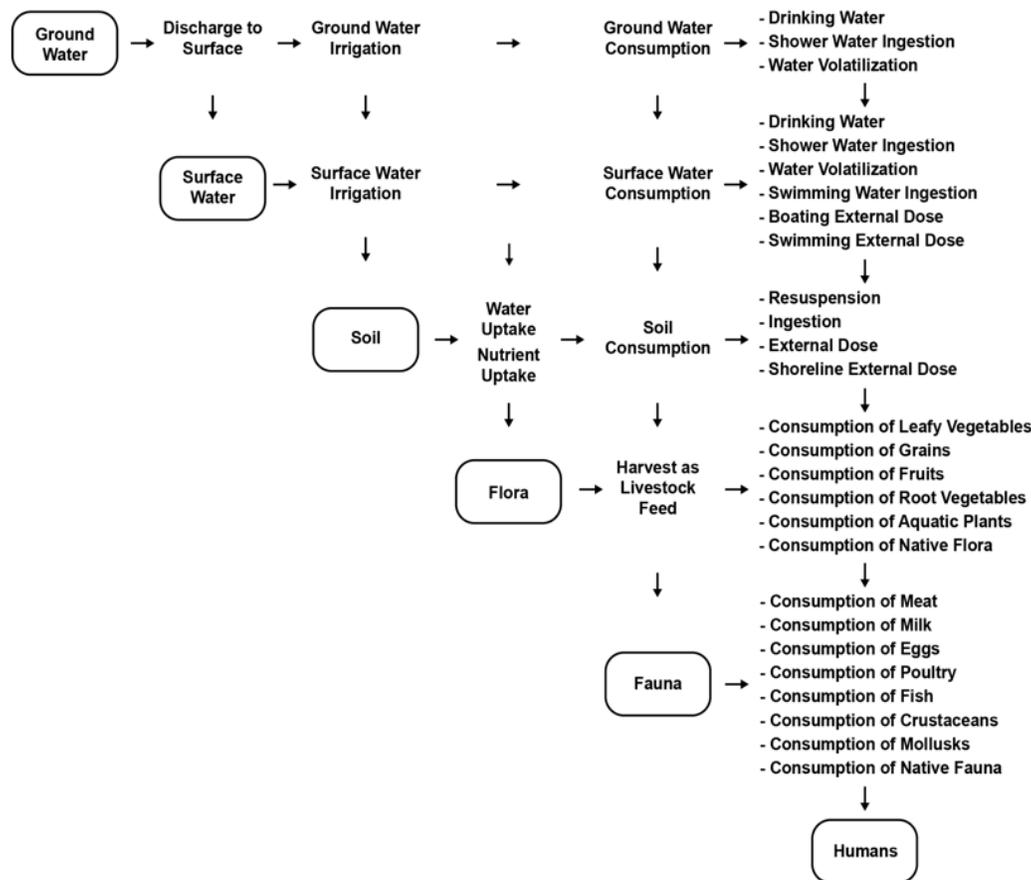


Figure 3-13 GDSA biosphere model interaction matrix (graphic from Condon et al. 2020)

3.2 Framework Development

In addition to development of the PFLOTRAN code (Section 3.1), there were advances in PA framework and PA methods around the code. These advances include development of:

- Next Generation Workflow graphical interface for GDSA Framework simulations (Section 3.2.1)
- Visualization and local grid refinement using VoroCrust (Section 3.2.2)
- Uncertainty and sensitivity analysis applications and methods (Section 3.2.3)
- International generic performance assessment models (Section 3.2.4)
- Geologic Framework Model application to crystalline rock (Section 3.2.5)
- Improved generic reference case models (Section 3.2.6)
- GDSA Calculation Archive (Section 3.2.7)

3.2.1 GDSA Framework Graphical Workflow

The GDSA computational framework is largely comprised of two primary computational capabilities: PFLOTRAN and Dakota, as described in Section 2.2. An analysis supporting the performance assessment of a geologic repository involves the development and use of many additional connective computational components beyond these two primary components, including the development of input files, scripts that

connect the capabilities of PFLOTRAN and Dakota, scripts to submit calculations to computational resources, capabilities to gather and post-process results, and much more. The collection of these computational components that comprise the complete body of work required to produce results of interest from a computational simulation capability, such as GDSA Framework, is hereafter referred to as an *analysis workflow*.

The development of analysis workflows for GDSA analyses commonly occurs on an analysis-by-analysis basis and requires the expertise of highly experienced modelers/analysts who are very familiar with the PFLOTRAN code base. These complex analysis workflows often involve many manual steps and continuous monitoring of simulations. This can make it difficult to replicate previous analyses, hand-off analyses between analysts, and/or train new analysts to produce analysis workflows as the reproducibility of these analyses depends on how well the modeler/analyst organized and documented what was done.

To improve analysis workflow automation, development, reproducibility, and traceability for repository PA simulations, the GDSA team began developing automated analysis workflows using the Next-Generation Workflow (NGW) capability in FY 2020. The NGW capability is an open source engine that was developed at Sandia National Laboratories to provide analysts with a capability to construct, execute, and communicate end-to-end computational simulation analysis workflows (Orient et al. 2020). This capability is a graphical, node-based interface that includes many pre-programmed support functions which are utilized within computational simulation analysis workflows. NGW is available within the Dakota Graphical User Interface (GUI) (Ridgway 2020) and is thus available to the GDSA analysis community.

Development of NGW analysis workflows for the GDSA Framework is focused on providing an automated workflow capability with the following objectives:

- Reduce the learning curve for new users to set up and run simulations and analysis workflows
- Speed up analysis workflow execution time
- Eliminate/reduce the need for manual intervention and allow for automated monitoring
- Reduce the potential for the introduction of human errors
- Increase traceability and reproducibility

A series of analysis workflows utilizing NGW for GDSA analysis exemplars, including a short course exercise and the Crystalline Reference Case, were developed in FY20 to pilot the use of this capability for the GDSA program as described by Mariner et al. (2020a). Section 3.2.1 of the FY20 GDSA Framework report (Mariner et al. 2020a) details the development of these workflows, provides examples of their use, and proposes a path forward for additional development. Appendix C of Mariner et al. (2020a) provides a complete graphical workflow, demonstrating an uncertainty quantification (UQ) study. This workflow sets parameters of interest, inserts these parameters into the PFLOTRAN input file, collects additional input files, runs PFLOTRAN, and develops and runs a post-processing script. The eventual goal of this work is to develop an analysis workflow library that can be made available to the GDSA analysis community.

The development and application of the analysis workflows for the Crystalline Reference Case described above continued in FY 2021 in support of the Uncertainty Analysis (UA) detailed in Swiler et al. (2021) and summarized in Section 3.2.3. The workflows developed for the UA are organized in a nested structure which is hereafter referred to as the *Crystalline Reference Case UA Nested Workflow*. Crystalline Reference Case UA Nested Workflow development, including the extension of the workflow to new quantities of interest (QoI) and hardening the workflow to be more robust to changes in systems

and supporting software versions, took place throughout the year to support the final UA production runs. Crystalline Reference Case UA Nested Workflow application took place throughout the year as the workflow development progressed and included testing as well as two full UQ studies for the Crystalline Reference Case. The benefits of the additional development and extended application of the Crystalline Reference Case UA Nested Workflow structure in FY 2021 included:

- Quicker turnaround for full UQ studies enabling two full studies in one year
- Easy handoff of workflows across analysts
- Increase in collaborative teamwork supporting final production runs enabled by version control and file sharing
- Increased robustness against HPC outages during production runs

A notional view of the detailed structure of the Crystalline Reference Case UA Nested Workflow is shown in Figure 3-14. A detailed description of each of the elements of this workflow along with an overall description of this structure is given in Swiler et al. (2021). An example of one of the elements is shown in Figure 3-15, which depicts the Local Resource Deterministic Workflow as it is implemented in NGW. This element sets up the input files, substitutes sampled parameter values, and sends all required inputs to a subsequent workflow that submits, runs, and post-processes the PFLOTRAN simulation on the HPC resource.

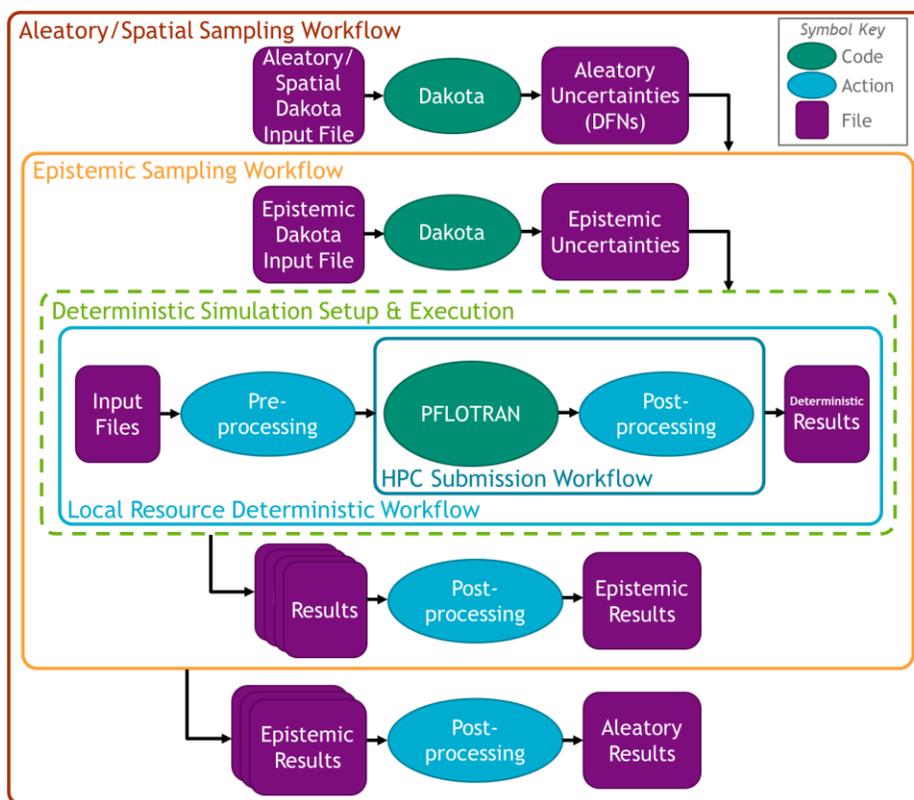


Figure 3-14 Detailed notional view of the Crystalline Reference Case UA Nested Workflow structure showing the workflows along with important files, actions, and codes that are employed for each.

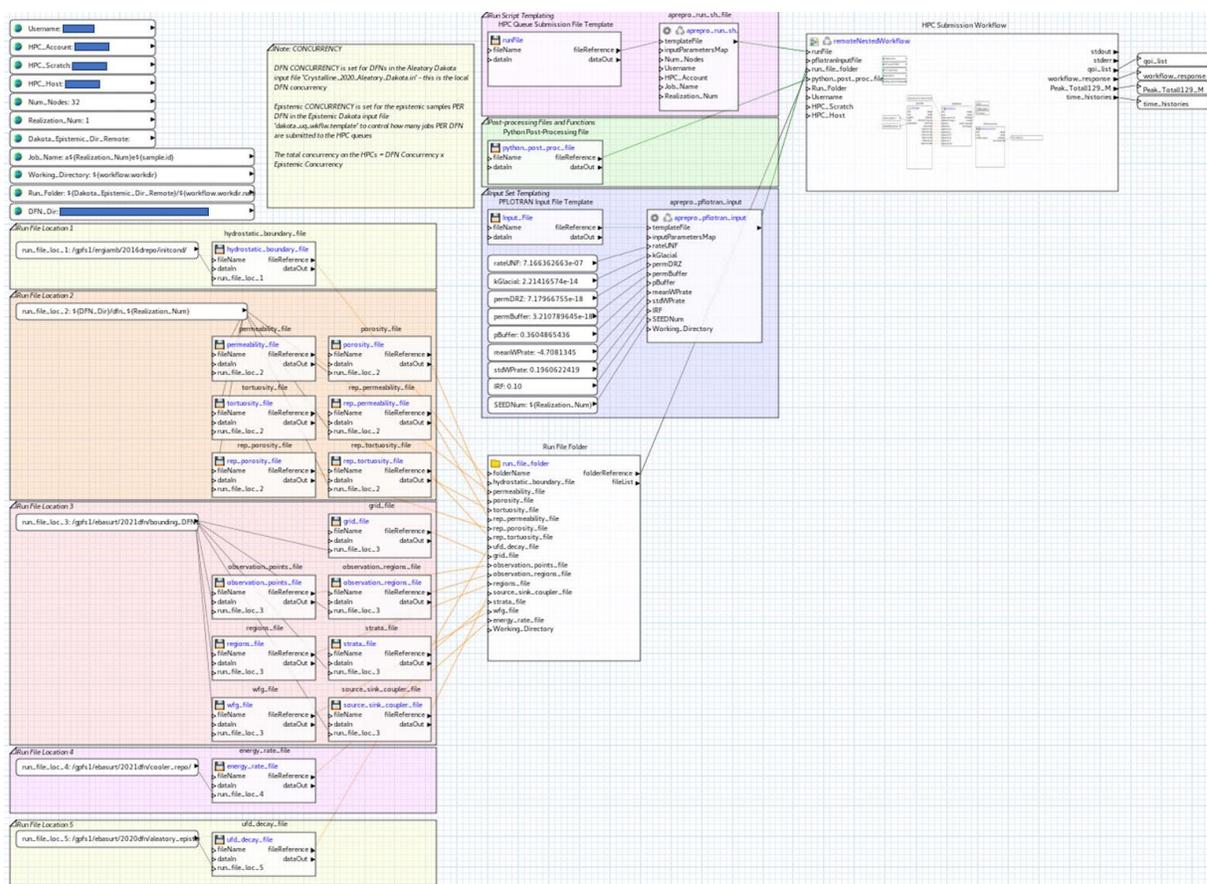


Figure 3-15 NGW implementation of the Local Resource Deterministic Workflow element of the Crystalline Reference Case UA Nested Workflow

In addition to the development and application of the Crystalline Reference Case UA Nested Workflow, an introductory training course for the NGW software and analysis workflows developed using this capability to support GDSA was developed in FY21 and was presented to GDSA code developers, analysts, and program management on March 8, 2021. This training covered the basic elements of the NGW software and included hands-on demonstrations to introduce attendees to using the software through the Dakota GUI. Finally, the training included a walk-through demonstration of the Crystalline Reference Case UA Nested Workflow structure described above. This training increased the awareness of the NGW capability across the GDSA team and promoted the use of this capability for future analyses beyond the Crystalline Reference Case UA. The training materials for this course are available upon request.

The workflow capabilities demonstrated for the Crystalline Reference Case will be expanded to additional studies and applications under the GDSA program in future work due to the successful application of these capabilities for the Crystalline Reference Case UA. In addition to extending this capability to new applications, documentation including instructions on setting up and running workflows will be developed to support the adoption of the workflow by new analysts. This will include recommendations for UQ studies and could leverage faster-running examples as study problems.

3.2.2 Voronoi meshing and simulation

Significant progress has been made on Voronoi meshing and PFLOTRAN simulations on the meshes this year. Two major advances in the VoroCrust meshing software have been achieved: Exodus output for visualization and local grid refinement. The Exodus II file format is necessary to create high-quality images of three-dimensional simulation results in ParaView. Local grid refinement allows users to have a fine simulation mesh in regions of interest and a coarse mesh in less important regions. This reduces the number of grid-cells and speeds up simulations.

LaForce et al. (2021) Section 6.2 shows mesh development for four test cases of common reservoir structures, along with qualitative and quantitative (where possible) error analysis for 16 simulation examples on the mesh. The simulation mesh on a slice through the interior of each volume is shown on Figure 3-16. The analysis indicates that the Voronoi meshes and simulations on them are accurate where analytical solutions exist, consistent where there are no quantitative benchmarks, and that results are reproducible across multiple realizations of the mesh.

A large-scale geological model of the Rock Springs uplift anticline has been meshed using VoroCrust and simulations have been run on them (LaForce et al. 2021, Section 6.3). This structure has been meshed in previous reports (LaForce et al. 2020, Sevougian et al. 2019c), but complex simulations were not possible without the new local grid refinement capability and the new PFLOTRAN solvers developed in Nole et al. (2021).

Utilizing local grid refinement, it is possible to define four evenly spaced gas injection wells and accurately simulate buoyant-gas flow beneath an interior geological material boundary. Four wells inject CO₂ at a combined rate of approximately 1 million tons per year for 100 years. As shown in Figure 3-17, the randomly-oriented grid is able to capture gas-fingering during the unstable gas displacement, which is difficult to do using structured meshes. Though the example is not directly applicable to GDSA workflow, this demonstrates that PFLOTRAN with VoroCrust meshes is capable of simulating highly complex problems using output from commercial geological mapping software at the large scale.

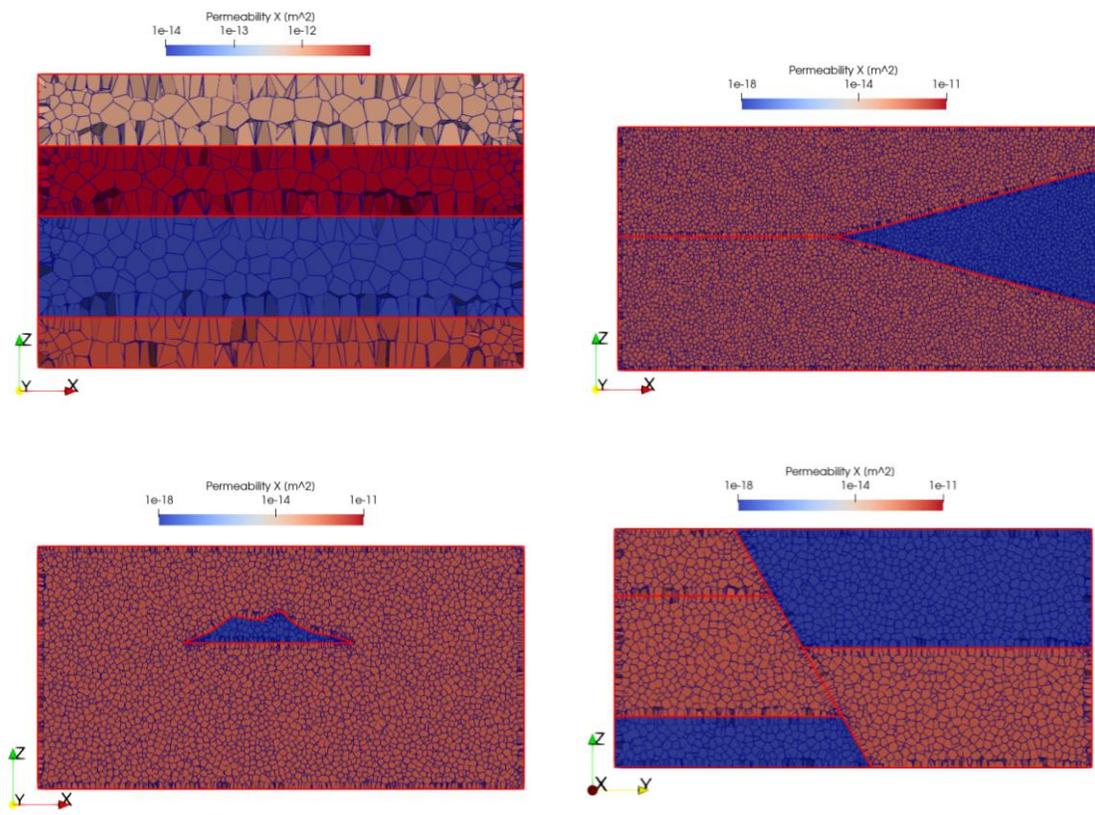


Figure 3-16 Simulation meshes for four test structure volumes showing permeability for heterogeneous simulations. Top Left: Test case 1 with 4 horizontal layers. Top Right: Test Case 2 with a pinch-out. Bottom Left: Test Case 3 with an interior lens. Bottom Right: Test Case 4 with layering and an offset fault. Test Cases 1-3 are shown on the XZ plane on a slice through the plane $y=10$, while solution for Test Case 4 is shown on the YZ plane on a slice through the plane $x=10$ (after LaForce et al. (2021)).

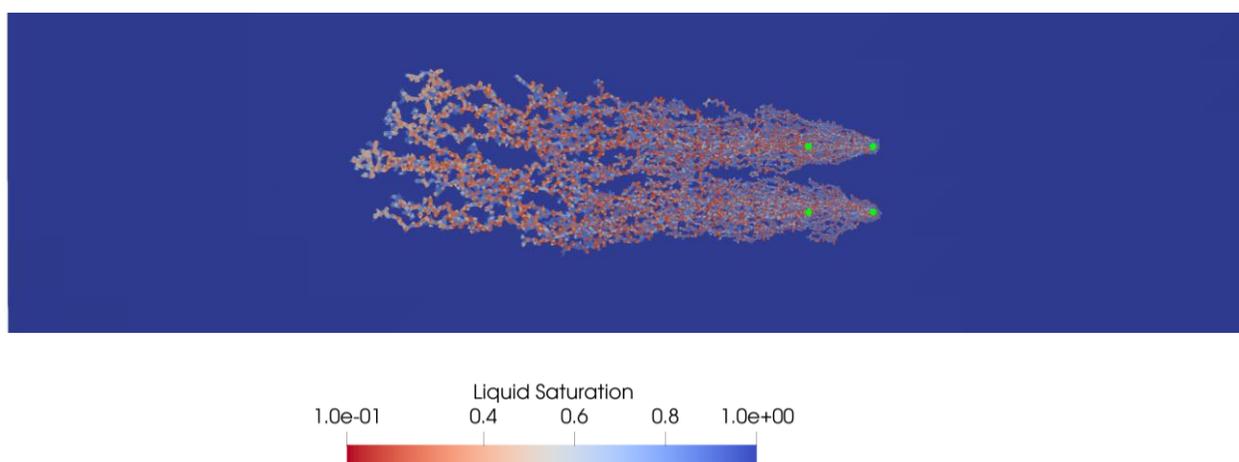


Figure 3-17 Liquid saturation at the top of the injection interval (plan view-i.e., x-y plane) at 5 years of injection. Green circles mark the well locations, which are 1 km apart.

3.2.3 Uncertainty and Sensitivity Analysis

The report *Uncertainty and Sensitivity Analysis Methods and Applications in GDSA Framework (FY2021)* (Swiler et al. 2021) describes the work done on UQ and sensitivity analysis (UQ/SA) in detail. This section provides a summary of that work.

This year, the GDSA UQ/SA team focused on three major areas:

- *Multifidelity UQ methods.* We demonstrated the use of multifidelity (MF) methods in the context of global sensitivity indices: they were used to calculate the Sobol' sensitivity indices which describe the fraction of an output variance that is attributable to the variance in an input parameter.
- *Discrete Fracture Network (DFN) analysis.* We performed more detailed analyses of the DFNs, including a study of the transmissivity relationships used in the DFN generation. We found that the correlated constant transmissivity gave much lower permeability fields than the depth-dependent transmissivity relationship, but that many of the QoI were not statistically significantly different when using a correlated constant vs. depth-dependent relationship.
- *Sensitivity analysis of an updated crystalline reference case.* Our SA study this year included an additional input parameter, instant release fraction (IRF), and several new QoI. We demonstrated a capability to calculate the sensitivity indices as a function of time to show how different input parameters may have a changing effect on a QoI over time (e.g., one parameter may be important in early years and another at 100K-1M years).

3.2.3.1 Multifidelity UQ Methods

The main idea in multifidelity UQ is to extract information from a limited number of high-fidelity model evaluations and complement them with a much larger number of a set of lower fidelity evaluations. The final result is an estimator of the mean response with a lower variance: a more accurate and reliable estimator can be obtained at lower computational cost. We presented results demonstrating the use of these multifidelity methods in calculating Sobol' sensitivity indices. The demonstration was done on a simplified crystalline reference case that retained many of the key qualities of the full reference case.

We examined the use of Polynomial Chaos Expansion (PCE) surrogates within the multifidelity framework. We found that the multifidelity framework for sensitivity analysis performed better for some QoI than others, depending on whether the QoI exhibited discrete/discontinuous behavior and whether spatial uncertainty was included. For the case with no spatial uncertainty, MF PCE was able to generate the global sensitivity indices for certain performance QoI (e.g., peak ^{129}I in the aquifer) at a fraction (2.5%) of the cost of calculating the sensitivity indices from a full set of high fidelity runs. Our study showed that a MF PCE which exploits a hierarchy of meshes has the potential to improve sensitivity analyses when evaluations at the finest level are limited. More work remains in the validation of the PCE models and the extension of the MF framework to the full production crystalline reference case.

An example of the Sobol' indices calculated with 828 high fidelity runs vs. an equivalent cost of 21.02 runs using multifidelity sampling (with a sample profile of 18 high fidelity complemented by 108 medium fidelity and 108 low fidelity) is shown in Figure 3-18. Note that both approaches produced nearly identical results for the following QoI: the peak ^{129}I concentration in the aquifer, the peak of the ratio of two fluxes (aquifer to east boundary and rock to east boundary), and the median residence time of a tracer in the repository. Examination of other QoI is ongoing.

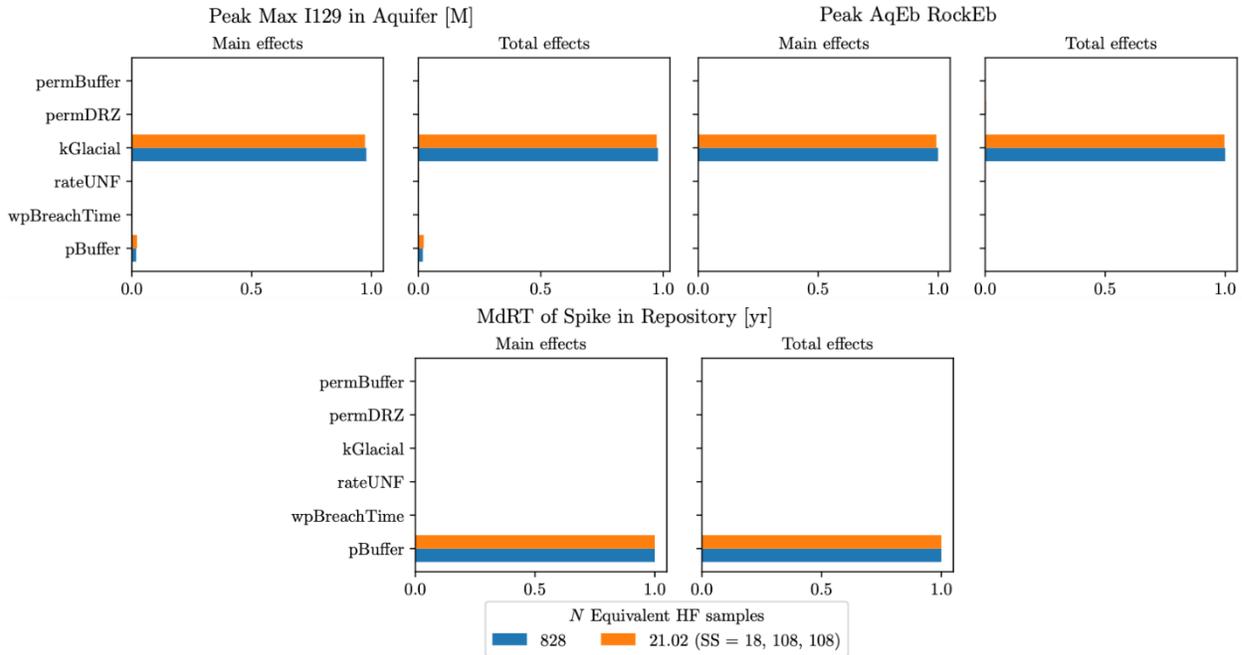


Figure 3-18 Comparison of Sobol' indices from the reference PCE computed with 828 high-fidelity samples (blue) and for the MF PCE (orange)

3.2.3.2 Discrete Fracture Network (DFN) Analysis

In FY 2021, members of the UQ/SA team used a recently added feature of dfnWorks to generate the DFNs for this analysis: correlated depth-dependent transmissivity. They examined the resulting averaged permeability fields in the ECPM and compared the impact of using correlated constant vs. correlated depth-dependent transmissivity.

The transmissivity approach had a large, statistically significant impact on the geometric mean of the permeability fields at each depth level in the repository. The two upper depth zones had higher averaged permeability fields when using the depth-dependent transmissivity.

The choice of transmissivity relationship did not have a major effect on some QoI such as peak ¹²⁹I or median residence time of a tracer in the repository. However, the transmissivity relationship chosen did show a significant effect on the water flux ratio (aquifer to east boundary flux / rock to east boundary flux). This is to be expected since the flow rate from the aquifer to the east boundary is likely minimally affected by the DFN while the flow rate from the rock to the east boundary was increased significantly for the depth dependent transmissivity relationship due to the large increase in permeability of the rock in depth zone 1.

This is a preliminary study of transmissivity but points to the significance of the DFN parameterization within the full repository performance assessment calculations. Figure 3-19 shows a boxplot of the peak ¹²⁹I in the aquifer for the two transmissivity relationships. Although the boxplots look a bit different, statistical tests indicate that we cannot say the mean values of the peak ¹²⁹I from these two sets of data are statistically significantly different.

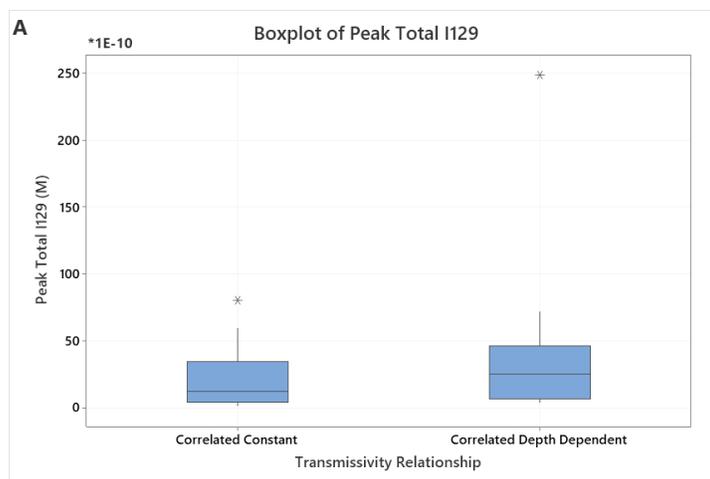


Figure 3-19 Boxplot of Peak Total ^{129}I based on 20 DFNs using a correlated constant transmissivity (left) vs. 20 DFNs with a correlated depth dependent transmissivity (right).

3.2.3.3 Sensitivity Analysis of an Improved Crystalline Reference Case

This year the GDSA UQ/SA team modified the repository in the crystalline reference case to make it cooler. Every other WP was removed and was replaced with buffer. As a result, WP temperatures reached a maximum of 120-130 °C, far below values reached previously. In addition, new DFNs were generated with the depth-dependent transmissivity relationship and new features were added, including soil compressibility. Also, the IRF was treated as an uncertain parameter. Finally, in the nested sampling approach, different epistemic realizations were generated for each spatial (DFN) realization.

Additional QoI measured this year include median residence time of a conservative tracer in the repository, fraction of a “spike” tracer remaining in the repository over time, and ratios of various water fluxes. These QoI were used to understand repository leakiness and retention and to distinguish characteristics attributable to flow through the different DFNs and the hydraulic connections between the repository and the aquifer.

Similar to last year, a sensitivity analysis was performed with and without the graph metrics included as parameters, where the graph metrics are proxy variables representing the spatial variability introduced by the DFNs. Details of the SA results both for scalar QoI as well as time-dependent QoI were presented (Swiler et al. 2021).

The time-dependent sensitivity analysis, where the Sobol’ sensitivity indices are calculated at many time points throughout the simulation, is a new capability added this year which involves a significant amount of computation and software infrastructure. For each time point, a surrogate must be constructed over all the simulation results for the QoI at that time point, and the SA calculations to determine the sensitivity indices must be performed. This additional computational infrastructure to support time-dependent SA has the benefit of more clearly showing how various parameters change their importance over time. With that information, we can better relate the sensitivity analysis results to the physics of the problem.

As an example of the time-dependent sensitivity indices and accompanying information generated from the PFLOTRAN runs, Figure 3-20 shows the sensitivity analysis results for the maximum ^{129}I concentration in the aquifer as a function of time.

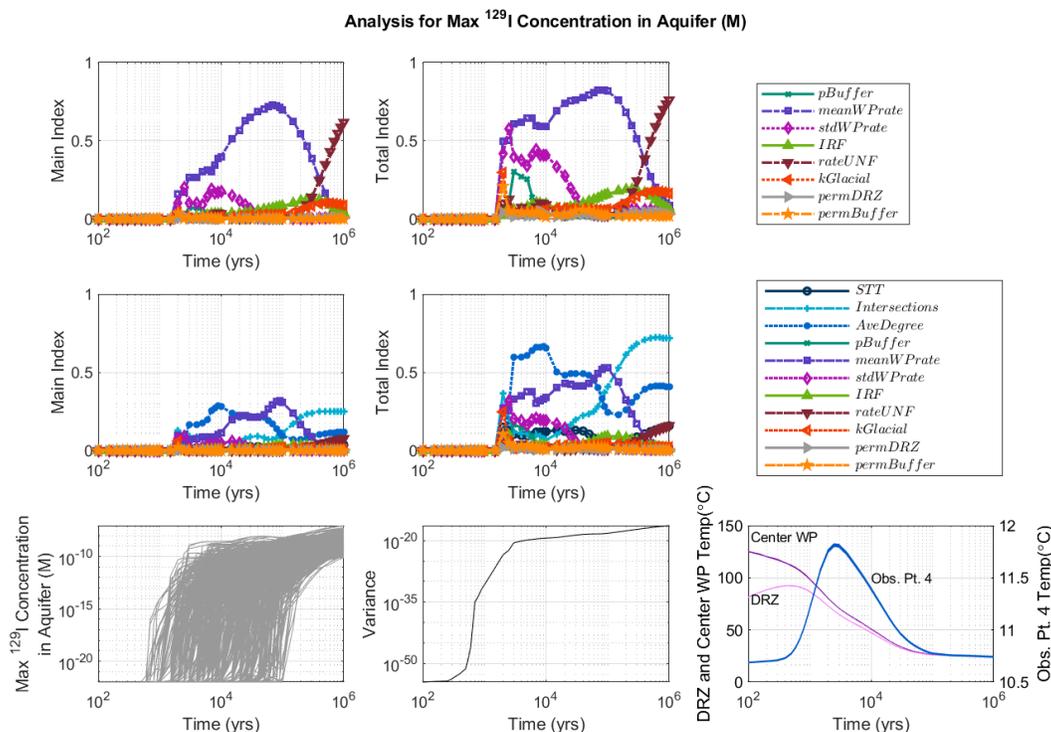


Figure 3-20 Global sensitivity analysis results over time for the maximum ^{129}I concentration in the aquifer

Specific findings from this year's sensitivity analysis study are listed below. The supporting analyses are documented in Chapter 4 of Swiler et al. (2021).

- Log-transformation improved results for ^{129}I . This transformation was justified by comparing the GSA results to scatter plots at different points in time to ensure that the relationships identified in the GSA are apparent in the data at corresponding times. This also made sense to do because of the very low variance in ^{129}I due to variation between low orders of magnitude.
- The uncertainty in IRF had less of an impact relative to the uncertainty of rateUNF on the peak ^{129}I concentration. This is likely because the uncertainty distribution on IRF was much narrower than that for rateUNF. However, based on the correlation between IRF and mean peak ^{129}I in the simulations, an IRF value of 0.1 generates nearly half of the ^{129}I in the mean peak ^{129}I concentration.
- Graph metrics tended to dominate sensitivity analysis results when included in the surrogate model construction, suggesting that spatial heterogeneity may drive uncertainty in the mean peak ^{129}I concentration in the aquifer in these simulations.
- Sensitivity analyses for fluxes, flux ratios, and repository retention of tracers made phenomenological sense. The addition of these QoI is an important advancement for the crystalline reference case.
- The number of fracture intersections with the repository correlates strongly with the direction of flow through the rock; there is more upward flow in the rock relative to horizontal flow as the number of these intersections increases. The number of these intersections also correlates with higher peak ^{129}I concentration in the aquifer, meaning fracture intersection avoidance has significant performance implications.

3.2.4 DECOVALEX-2023 Task F

The DECOVALEX project is an international research and model comparison collaboration for advancing the understanding and modeling of coupled THMC processes in geological systems (decovallex.org). Task F of DECOVALEX-2023 (LaForce et al. 2021) is a task that focuses on comparison of models and methods used for post-closure PA. Members of the GDSA team at SNL are leading this effort. The goal of this work is to test and build confidence in the models, methods, and software used for post-closure PA and to identify additional research and development needed to improve PA methodologies.

In Task F, two hypothetical repositories are being developed, one in crystalline rock and the other in salt. In the first year of the four-year task, nine teams from six countries participated in the crystalline repository and benchmarking exercises, and three teams from three countries defined a generic salt repository reference case. In the second year, each focus group gained one additional team.

To date, Task F has provided and will continue to provide numerous opportunities for learning new modeling approaches, developing new models for use in PA simulations, testing uncertainty and sensitivity analysis methods, comparing PA methods, and exchanging ideas with modelers in other programs. Several accomplishments in the past year are highlighted in the subsections below. Additional information on the status of this work is documented in LaForce et al. (2021).

3.2.4.1 *Crystalline*

The primary focus of the crystalline group during the first year of Task F was the testing and comparison of codes used to simulate flow and transport through fractured rock. Several benchmark cases were simulated by participating teams. This exercise allowed participants to examine differences in model implementation, types of model outputs, and the influence of modeling choices.

One of the exercises the SNL team took was to simulate the analytical solution of Tang et al. (1981) discussed in Section 3.1.3.4 and shown earlier in Figure 3-7. This exercise confirmed that PFLOTRAN can simulate with high accuracy the effects of fracture flow coupled with matrix diffusion in an equivalent continuous porous medium (ECPM).

The primary benchmark exercise of the Task F teams during FY 2021 was a 4-fracture DFN problem with and without stochastic fractures. Figure 3-21 shows the Uniform ECPM representation of the four deterministic fractures and the permeabilities of the cells affected by those fractures. Not all teams used an ECPM, several teams applied multiple models, and two teams applied particle tracking. In general, certain models performed better than others.

Breakthrough curves generated by the SNL team for the 4-fracture DFN problem without stochastic fractures are shown in Figure 3-22. Conservative tracer is introduced into the fracture at the left boundary, and the breakthrough curves are measured at the right boundary. Three of the breakthrough curves in the figure are from simulations that use PFLOTRAN (and the advection-dispersion equation, ADE) and one is from a simulation that uses particle tracking. Two use the DFN directly and two use an ECPM generated from the DFN. The Octree ECPM, which allows for a tetrahedral mesh and is refined only where there are fractures, resulted in a total of 30,086 active cells. The Uniform ECPM for this problem uses 9704 active hexahedral cells (Figure 3-21) (LaForce et al. 2021). In the Octree ECPM, all cells in the domain are active, whereas in the Uniform ECPM only cells simulating fracture behavior are active.

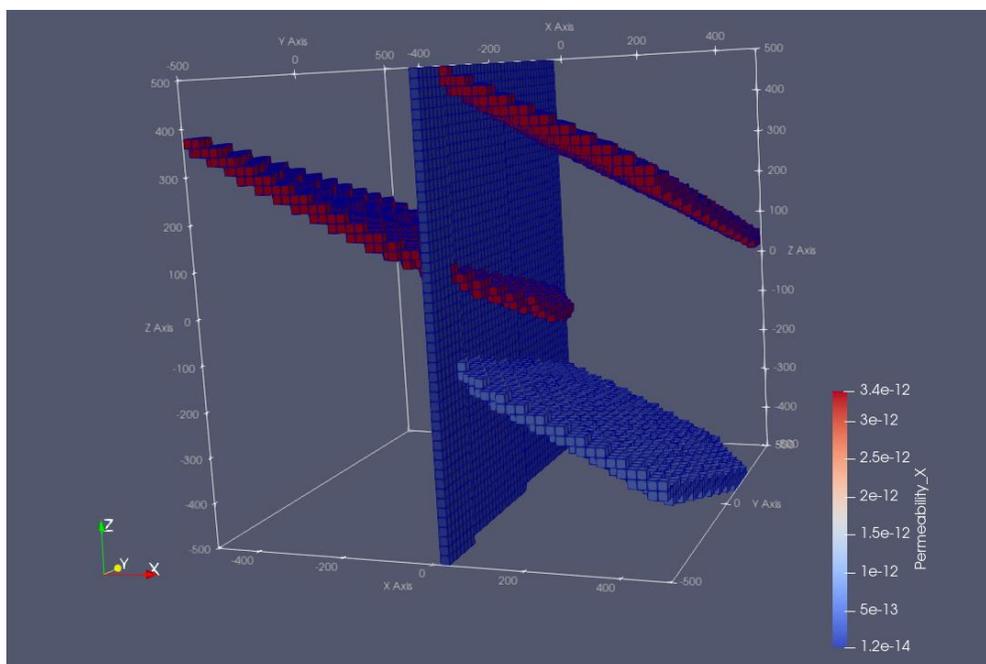


Figure 3-21 Uniform ECMP representation of the four deterministic fractures of 4-fracture DFN problem

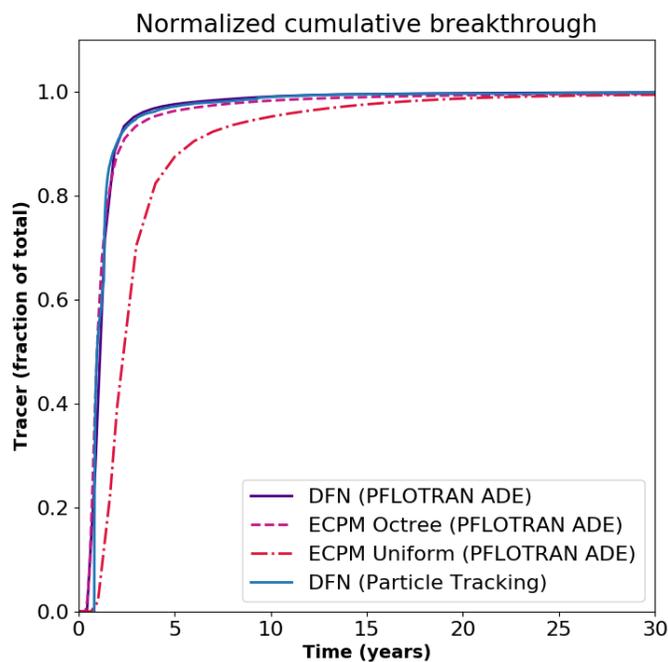


Figure 3-22 Breakthrough curves of SNL simulations of the 4-fracture problem without stochastic fractures. Three simulations use PFLOTRAN and the advection-dispersion equation (ADE). Two use the DFN directly while two use an ECMP generated from DFN mapping. The ECMP Octree mesh (30,087 active cells) is much more refined at fracture locations than the ECMP Uniform mesh (9704 active cells).

Figure 3-23 shows the 4-fracture system with the addition of families of stochastic fractures. Simulation by the various teams of the 4-fracture problem with stochastic fractures is ongoing.

The benchmarking of fracture and fracture network problems will continue, but the new focus of Task F is the development of a common crystalline repository reference case for each team to simulate. Many choices will be made regarding repository layout, boundary conditions, engineered barriers, and performance metrics as the teams prepare to simulate the new generic reference case.

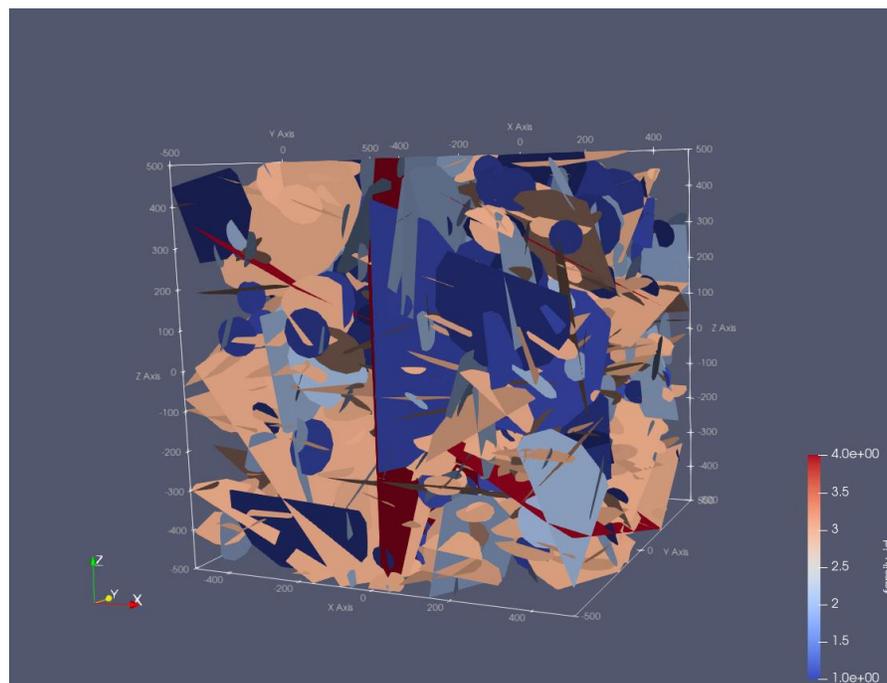


Figure 3-23 4-fracture DFN with stochastic fractures

3.2.4.2 Salt

In FY 2021, the salt group of DECOVALEX-2023 Task F developed a common reference case for simulation. The general setting is a generic salt dome as shown in Figure 3-24. The repository is located at a depth of 850 m and is accessed by two vertical shafts. A plan view of the repository layout is shown in Figure 3-25. The spacing of waste was chosen to ensure that temperatures on the waste package surface will not exceed 100°C.

The teams will simulate a scenario in which shaft seals fail at 1000 years allowing brine to flow down the shafts and into the repository. In addition to the information shown in Figure 3-24 and Figure 3-25, the teams have agreed upon:

- SNF and HLW casks, canisters, and inventories
- Cask and canister heat outputs over time
- Waste form IRFs and degradation models
- Vertical emplacement boreholes, backfills, and drift seals
- Shaft seal design
- Material properties

- Geochemical environment
- Initial conditions
- Outputs for comparison

These details are presented and discussed in Section 2 of LaForce et al. (2021).

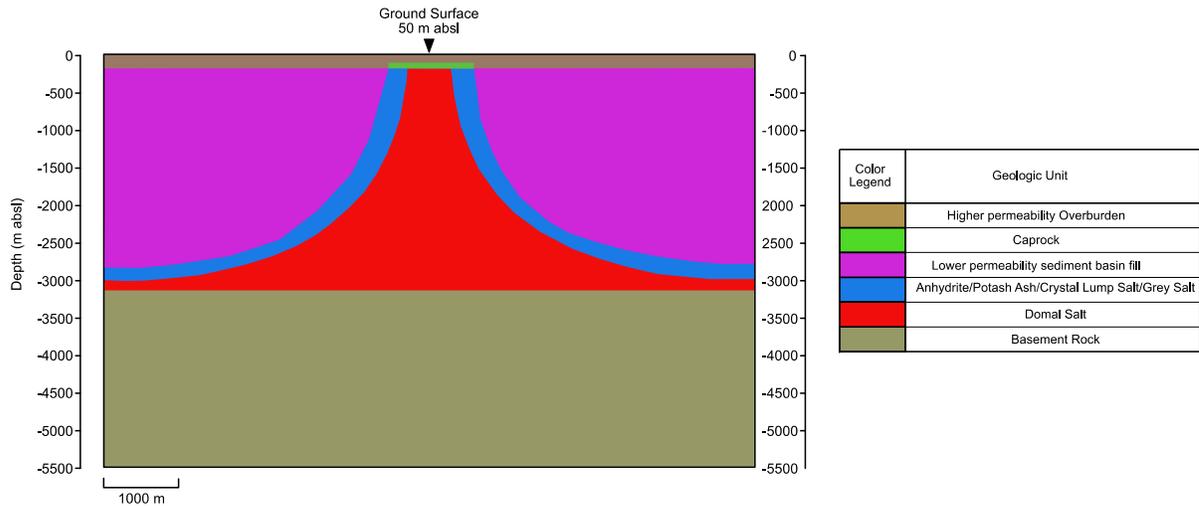


Figure 3-24 Geologic cross-section of salt reference case for Task F

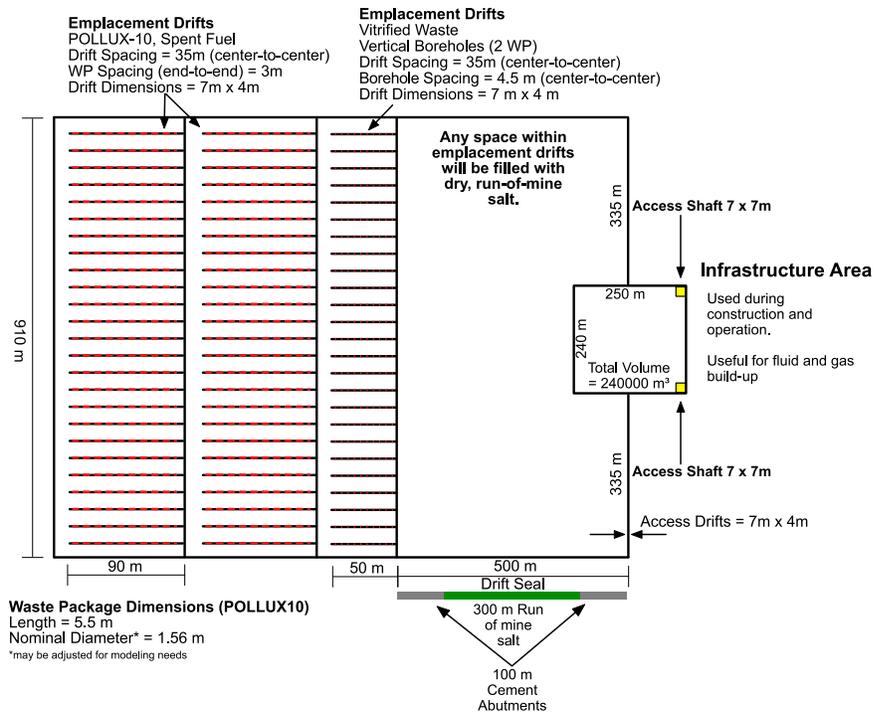


Figure 3-25 Repository layout for generic salt reference case for Task F

3.2.5 Geologic Framework Model

Host-rock models for the crystalline reference case have focused on a generic representation of the geosphere based partly on features of crystalline rock (e.g., fracture systems) that have been well documented at the Forsmark site in Sweden (Mariner et al. 2016). In FY 2021, we have developed a conceptual model of the shallow geosphere (within ~100 meters of the surface) based on features characteristic of the crystalline environment in the northeastern U.S. (LaForce et al. 2021). The primary features of the model are horizontal transmissive fractures in the uppermost crystalline bedrock and several types of glacial deposits of differing permeability that overly bedrock. Glacial deposits represented in the conceptual model are glacial till, glaciofluvial (“outwash”) sand and gravel deposits and glaciomarine silts and clays. These three glacial deposits are considered to be the most significant in terms of areal distribution, near-surface hydrology and flow and transport in surficial deposits (Thompson 2015). Data to support the conceptual model are described in detail in LaForce et al. (2021), with a short overview provided below.

3.2.5.1 Data to Support the Conceptual Model

Areas in New England were used to form the physical basis of the conceptual model. These areas are characterized by low topographic relief, a stable tectonic setting and a Quaternary history of glaciation resulting in thin glacial deposits that overlie the crystalline bedrock (LaForce et al. 2021). These characteristics are similar to those at the Forsmark site, allowing us to identify similarities and differences in the two settings that might inform how flow and transport is modeled in the shallow geosphere.

The area around the Sebago Lake Pluton of Maine is an example of one of the reference areas (Figure 3-26). Glacial till is the oldest glacial deposit and covers 80-90% of the landscape. Glacial outwash deposits composed of sand and gravel form isolated, elongated deposits that fill valleys and paleo-valleys. Regionally, these deposits cover about 4 percent of the landscape. Due to their high permeability, sand and gravel deposits are the major aquifers in the region. Because their deposition was largely confined to channels in valleys and broad lowlands, their distribution in the region is highly localized (Figure 3-26). This is potentially significant to both transport and biosphere models in that the most permeable deposits (and highest producing aquifers) occupy relatively small and discontinuous areas of the overall region. Fine-grained glaciomarine silt and clay deposits are present on the eastern margin of the reference area. Where present, silt and clay deposits form the major confining unit within the region.

Near-surface horizontal fractures/joints are a common feature in the region and are noted as significant transmissive features throughout southern Maine (e.g., Figure 3-27; Hansen et al. 1999). These fractures have been observed in quarries at depths as great as 40 and 80 meters (Dale, 1923). Joint spacings of 15-20 meters have been observed bounding the deepest sheets (Dale 1923). Horizontal joints generally bound lenticular sheets that gradually taper near their ends at joint intersections. Individual fractures extend for lengths of at least several tens of meters. Analogous horizontal fractures at Forsmark are interpreted to be hydraulically connected over distances of 2-3 km (Follin et al. 2008, Johansson 2008).

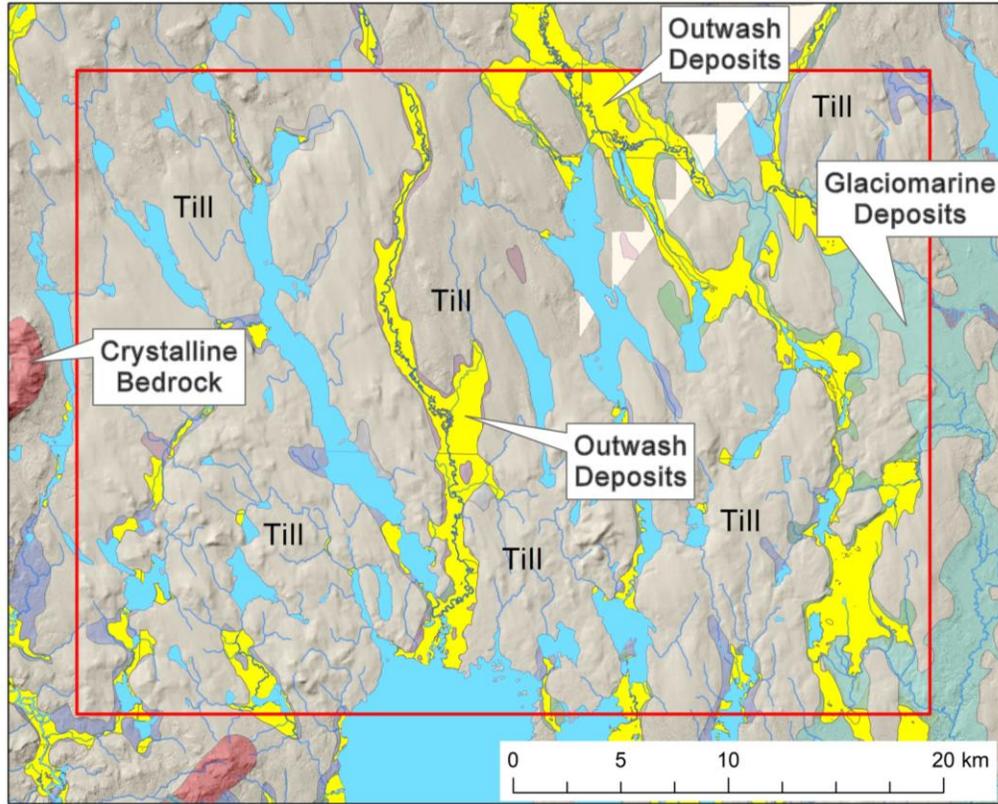


Figure 3-26 Surficial geology of the Sebago Lake Pluton reference area



Figure 3-27 Example of groundwater flow (icicles) in a horizontal fracture in a shallow crystalline rock exposure in southern Maine (MGS 2017)

3.2.5.2 Conceptual Model of the Shallow Geosphere

Based on interpretation of features in the shallow geosphere, a conceptual model of features and stratigraphy can be constructed and depicted as a schematic cross-section (Figure 3-28). The cross-section includes the boundary of marine submergence to show differences of deposits in marine and terrestrial environments. The distribution and properties of glacial deposits play a significant role in groundwater flow and transport and form the interface between the crystalline basement and the biosphere. Till deposits with permeability of $\sim 10^{-14} \text{ m}^2$ and a thickness of 2-3 meters can be assumed to be continuous over the entire landscape. Interspersed on the landscape at intervals on the order of 10 km are highly permeable (10^{-11} m^2) glaciofluvial sand and gravel deposits that generally coincide with modern drainage channels. In some areas, sand and gravel was deposited on top of marine silt and clay (Thompson 1987). Sand and gravel deposits allow enhanced localized lateral groundwater flow in the surface environment, ultimately discharging into streams and lakes. Radionuclides are expected to be transported advectively and diluted as they move horizontally through the sand and gravel deposits. In areas below the marine limit, marine silt and clay deposits act as confining units that overlie either till or sand and gravel aquifers. A summary of the properties of the glacial deposits is presented in Table 3-4.

Crystalline bedrock depicted below the surficial deposits includes a shallow zone of horizontal, highly transmissive fractures. Based on analogous fractures at Forsmark (e.g., Follin et al. 2008), horizontal sheet fractures are expected to have a significant role in near-surface groundwater flow. Hydraulic properties of the shallow horizontal fracture system are only known through studies of analogous features at Forsmark (Follin et al. 2008; Johansson 2008). Permeability of the fractured zone is in the range of 10^{-13} to 10^{-10} m^2 (Johansson 2008). Shallow horizontal fractures extend to a depth of $\sim 100 - 150$ meters at Forsmark and have been observed at a depth of 80 meters in quarry exposures in Maine.

Shallow horizontal fractures/sheet joints, believed to be analogous to those observed in Maine, have a significant role in conceptual models of shallow groundwater flow at Forsmark (Follin et al. 2008; Johansson 2008; Berglund et al. 2013). As understood from studies at Forsmark, transmissive horizontal fractures act as a groundwater drain that captures deep groundwater from below and recharge from above. As summarized in Johansson (2008) and Follin et al. (2008), horizontal fracture zones (corresponding to fracture domain FFM02 at Forsmark) are thought to be hydraulically connected over distances of 2-3 kilometers. When groundwater from depth intersects the shallow horizontal fractures, it is impeded from discharging locally and is instead carried laterally to discharge at the edges of the horizontal fracture domain (Johansson 2008; Berglund et al. 2013). The findings at Forsmark suggest that the shallow horizontal fractures are not present uniformly beneath the landscape but are instead grouped into domains with an extent of a few kilometers (Follin et al. 2008).

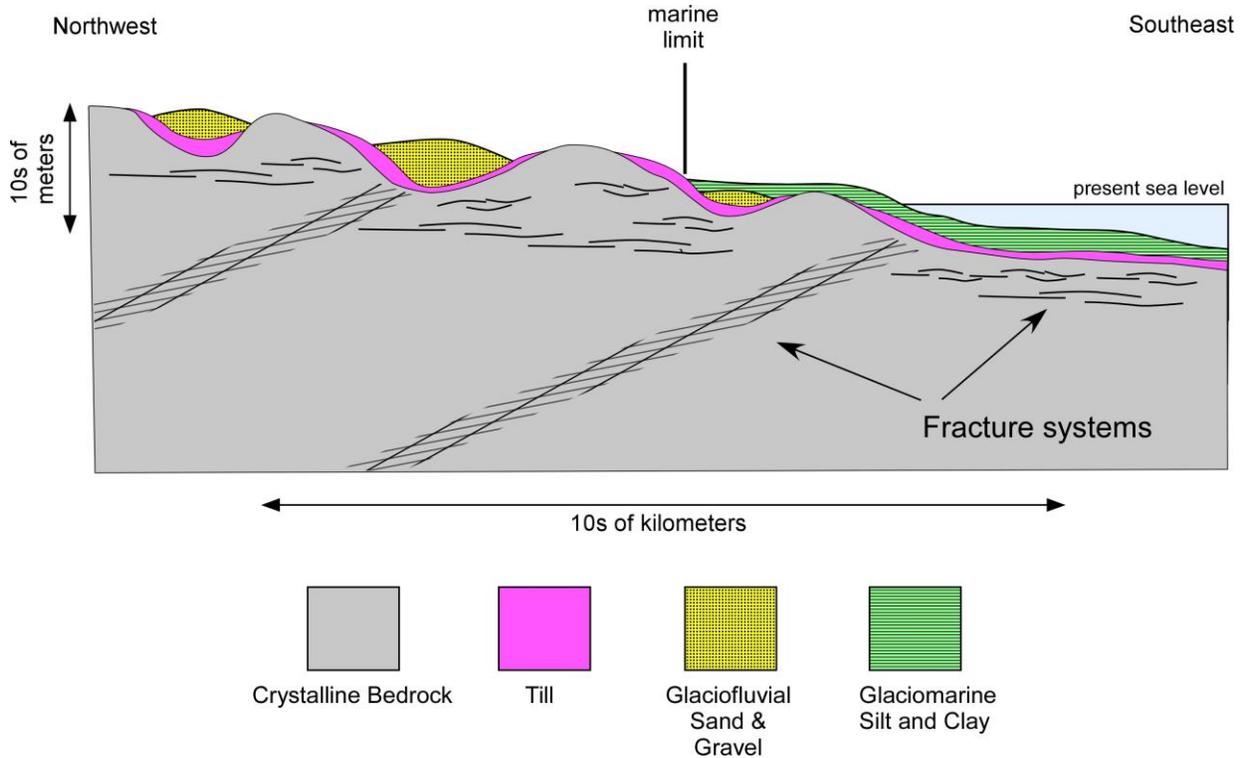


Figure 3-28 Conceptual model of stratigraphy and fracture systems of the shallow geosphere. Vertical scale is highly exaggerated

Table 3-4 Representative values of properties for glacial deposits included in the conceptual model

| Property | Till | Glaciofluvial sand and gravel | Marine silt and clay |
|--------------------------------|--|--|--|
| Typical thickness (m) | 2 – 3 | 5 – 25 | 10 – 40 |
| Permeability (m ²) | 10 ⁻¹⁴ – 10 ⁻¹³ m ² | 10 ⁻¹¹ – 10 ⁻¹⁰ m ² | 10 ⁻¹⁷ – 10 ⁻¹⁵ m ² |
| Porosity | 0.2 – 0.4 | 0.25 – 0.45 | 0.4 |
| Typical well yield (gpm) | 1 – 2 | 10 – 1000 | Not available |

3.2.6 Repository Reference Cases

Over the past decade, an assortment of generic repository reference cases has been developed and simulated using PFLOTRAN and GDSA Framework. Table 3-5 identifies the core reference cases and their associated documentation. They include reference cases in different host rocks and different designs. New developments in these reference cases in the past year are highlighted in the subsections that follow.

Table 3-5 Repository concepts and generic reference cases implemented with GDSA Framework

| Repository Type(s) | Conceptual Models | Recent Model Applications |
|--|--|---|
| Argillite/shale repository | | |
| SNF ranging from 4-PWR waste packages to 37-PWR DPCs | Jové Colón et al. (2014); Zheng et al. (2014) | Mariner et al. (2017); Sevougian et al. (2019) |
| High-temperature shale repository | Stein et al. (2020) | Stein et al. (2020) |
| Crystalline repository | | |
| Commercial SNF | Wang et al. (2014) | Mariner et al. (2016); Swiler et al. (2019, 2020, 2021) |
| DOE managed waste (cancelled by DOE in 2017) | | Sevougian et al. (2016) |
| Salt repository | | |
| Commercial SNF | Sevougian et al. (2012); Freeze et al. (2013) | Sevougian et al. (2016); LaForce et al. (2020) |
| DOE managed waste (cancelled by DOE in 2017) | | Sevougian et al. (2019) |
| Alluvium repository, unsaturated conditions | | |
| SNF ranging from 12-PWR waste packages to 37-PWR DPCs | Mariner et al. (2018) | Mariner et al. (2018); Sevougian et al. (2019); LaForce et al. (2021) |
| Dual purpose canister (DPC) | | |
| 24- to 37-PWR DPCs and 68- to 80-BWR DPCs | Price et al. (2019a) | Price et al. (2019b) |
| Deep borehole disposal (cancelled by DOE in 2017) | | |
| Various waste types, including Cs/Sr capsules | Brady et al. (2009) | Freeze et al. (2016, 2019) |

3.2.6.1 Shale

Early in FY 2021, new conceptual models for a high-temperature shale reference case were defined. Two types of shale were distinguished, ductile and brittle, and multiple options were presented for repository layout and engineered barriers. Full details of these designs are documented in Stein et al. (2020).

The results of new simulations of the previously developed lower-temperature shale reference case are shown in Section 3.1.3.3. Those simulations were used to demonstrate the implementation of two FMD model surrogate models that were updated in FY 2021. The results indicate that FMD is more rapid than using the traditional fractional degradation rate model for the applied conditions.

In addition to the full reference case, a smaller scale shale submodel was developed and simulated in FY 2021. The primary purpose of this smaller scale model is to examine the effects of resaturation and swelling on buffer and DRZ evolution. A diagram of the model domain is provided in Figure 3-29. Results to date indicate that the reduced-order buffer swelling model implemented in PFLOTRAN generally simulates the geomechanical behaviors predicted by the fully-coupled mechanical models. However, there is still more work to be done. A full analysis of this exercise and projections for future work are provided in LaForce et al. (2021).

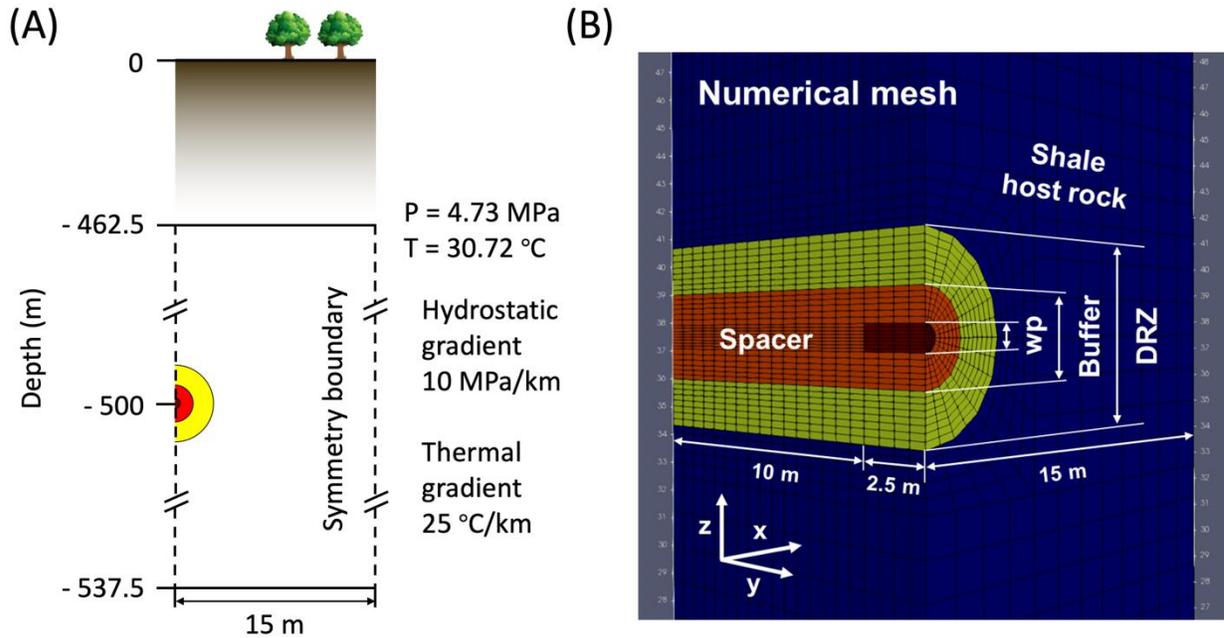


Figure 3-29 Shale repository submodel (A) and mesh (B)

3.2.6.2 Crystalline

Advances in the crystalline reference case in the past year were in the reference case model itself and in the conceptual geologic framework model of the near surface and overburden. The advances in the implemented model are documented in parts of Swiler et al. (2021). Those for the near surface are documented in Section 3.2.5 above and in LaForce et al. (2021). In FY 2022, a plan will be developed to implement the identified near surface features and characteristics into the crystalline reference case so that it can be connected to the new biosphere model under development (Section 3.1.3.8).

A diagram of the current crystalline reference case model domain and implementation is shown in Figure 3-30. Important adjustments and additions to the crystalline reference case simulations this year include:

- A cooler repository was generated by removing every other WP (maximum temperature of the centermost WP decreased from 200°C to 130°C)
- An uncertainty distribution for the IRF was implemented
- Additional QoI were automated in the workflow to measure median residence time of initial tracer in the repository and other QoI for DFN realizations and sensitivity analysis
- A new sampling scheme was implemented to increase the sampling of epistemic uncertainty

The resulting simulations indicate that an IRF of 0.1 for ^{129}I accounts for nearly half of the peak ^{129}I concentration in the aquifer. They also indicate that the number of fracture intersections with the repository correlates with upward flow through the rock and higher peak ^{129}I concentrations (Swiler et al. 2021). Additional analysis is planned for examining relationships between repository retention time, water flux ratios, and peak ^{129}I concentrations.

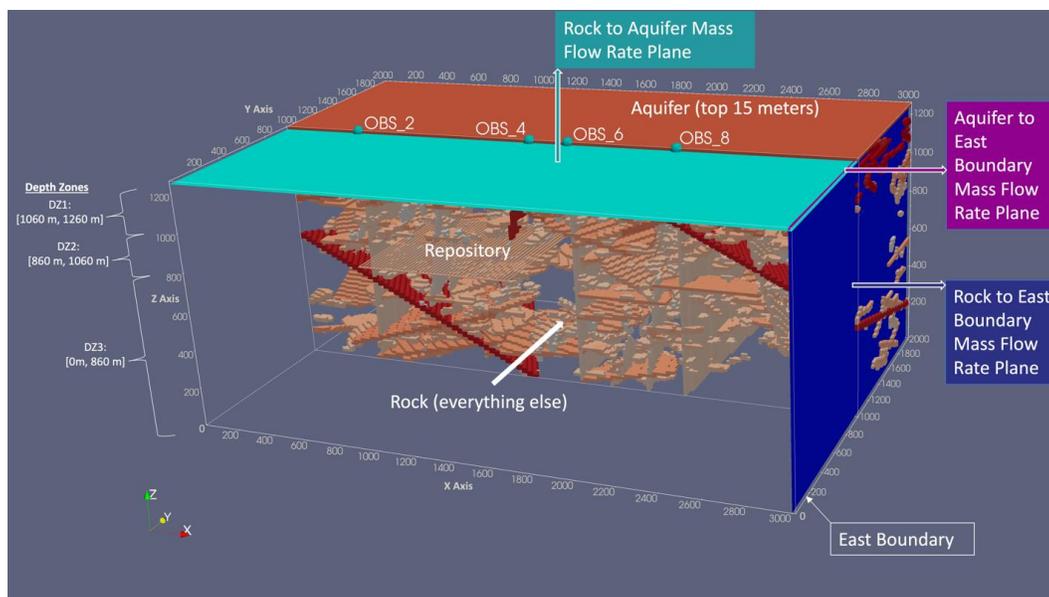


Figure 3-30 Crystalline reference case ECPM realization showing the repository, a midline cutaway of the cells propagating fracture properties, and the surfaces where water flow QoI are measured

3.2.6.3 Salt

GDSA activity in the salt reference case in FY 2022 focused on the development of a new conceptual model and design for a salt dome reference case. That work aligned with DECOVALEX-2023 Task F. Section 3.2.4.2 above highlights the related accomplishments on that model this year. Details of that work are documented in LaForce et al. (2021).

3.2.6.4 Unsaturated Alluvium

As documented in Section 5 of LaForce et al. (2021), there were several important advances regarding the unsaturated alluvium reference case in FY 2021. Figure 3-31 shows the general lithology and repository layout implemented. Details are provided in LaForce et al. (2021).

The new smoothed characteristic curves (Section 3.1.2.1) and improved PFLOTTRAN solvers (Section 3.1.2.2) allowed, for the first time, complete simulation of this reference case. Two infiltration rates were simulated, 1 and 10 mm yr⁻¹. Example results are shown in Figure 3-32. As indicated, dry conditions are calculated to occur early in the simulation and to continue for hundreds of years.

In addition, a smaller-scale model was developed and simulated to study the effects of meshing for a single DPC in the near field of an unsaturated alluvium repository. The results, shown in Figure 3-33, indicate that the Voronoi mesh approach produces more realistic contours of gas pressure and gas saturation than could be obtained from an unstructured grid approach (LaForce et al. 2021).

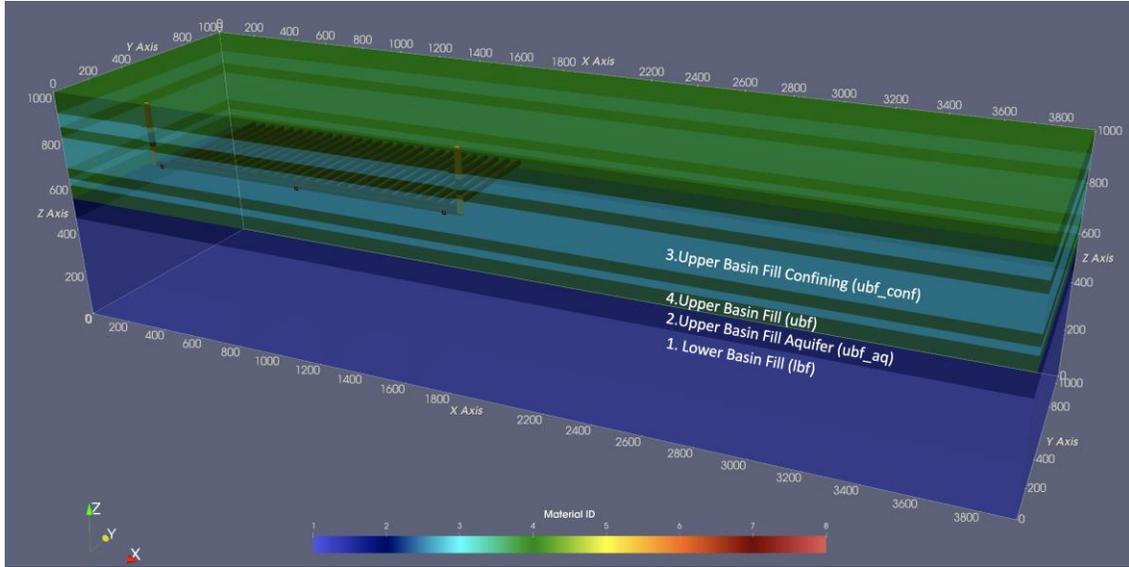


Figure 3-31 Half-symmetry model domain of the unsaturated alluvium reference case (3915 m × 1065 m × 1005 m). The repository is located at a depth of 250 m. The half-symmetry repository includes 27 drifts containing 25 WPs (24-PWR, 100 y OoR, 40 GWd/MTU burn-up) spaced 20 m apart center-to-center. Drift spacing is 50 m.

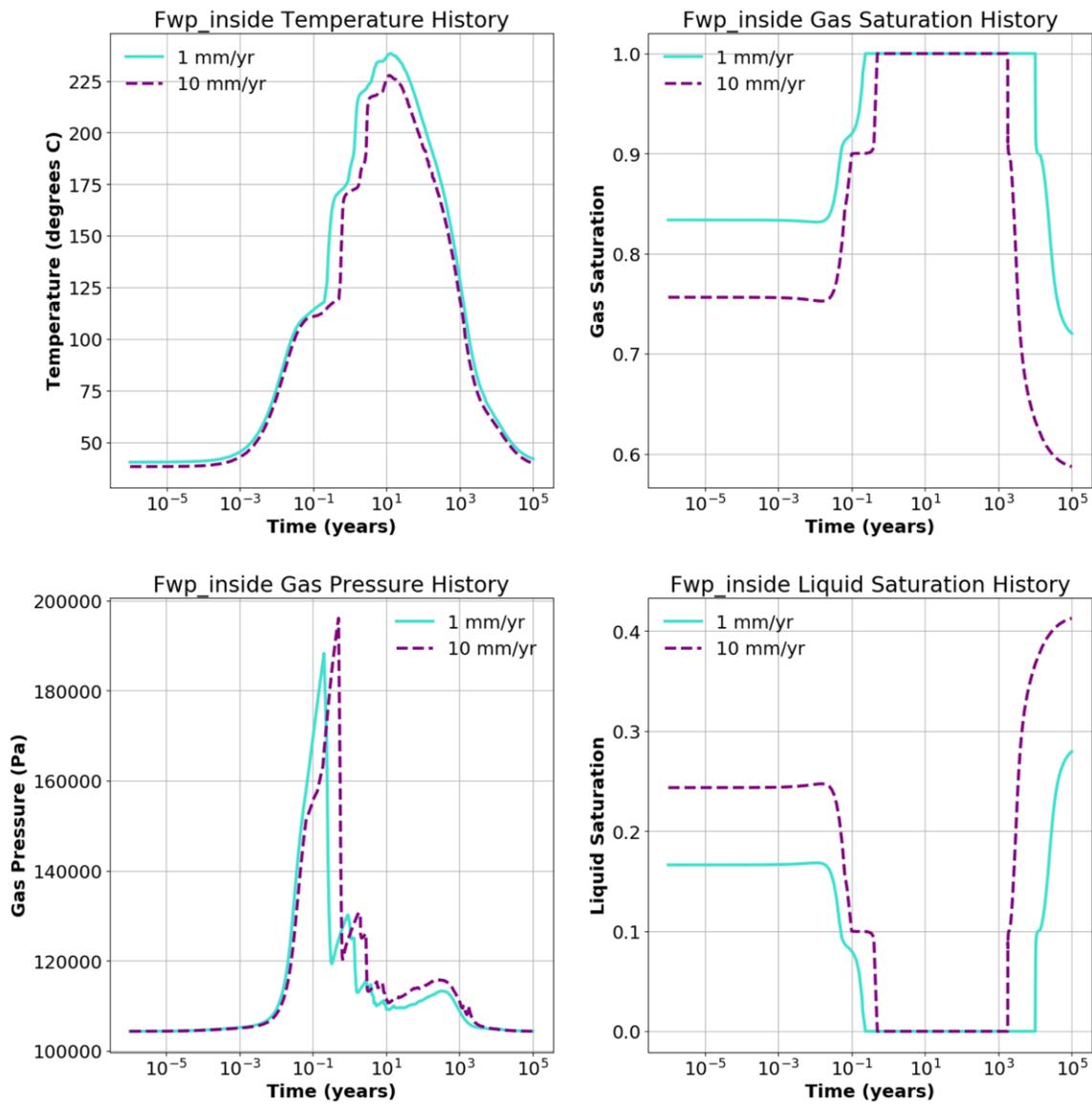


Figure 3-32 Temperature, gas saturation, gas pressure and liquid saturation time histories at the “Fwp_inside” location (centermost WP) in the repository

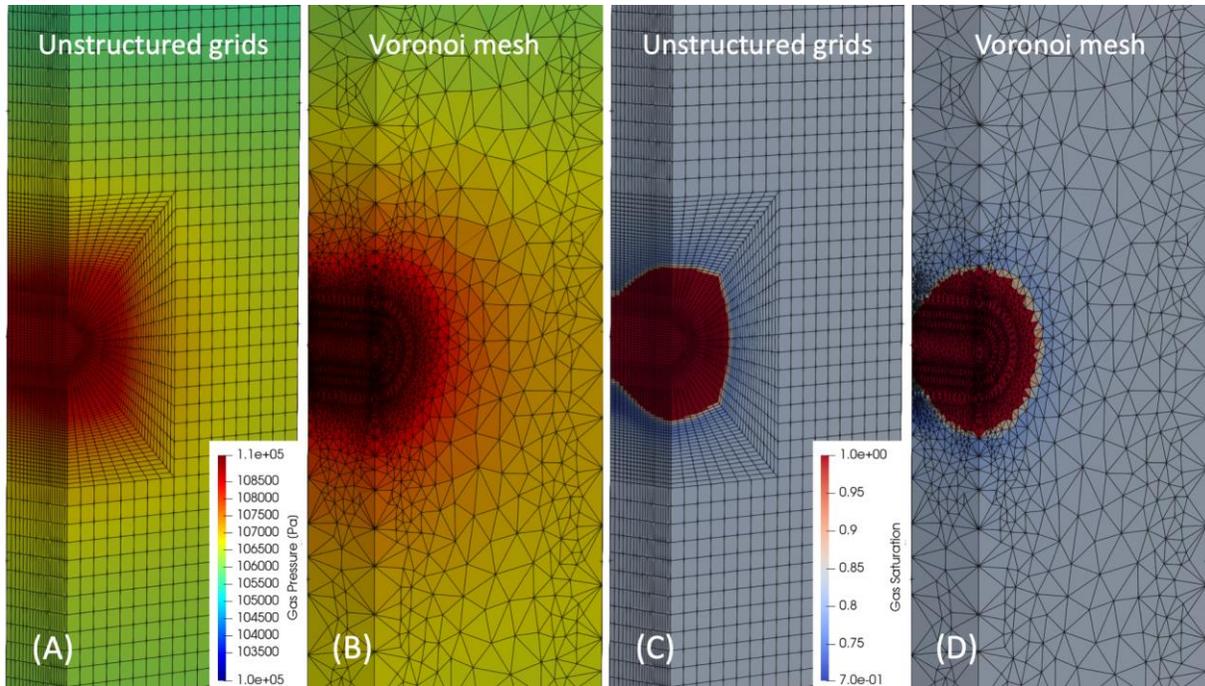


Figure 3-33 Gas pressure (A, B) and gas saturation (C, D) results from PFLOTRAN models using unstructured grids and Voronoi meshes at $t = 100$ years for an unsaturated alluvium repository

3.2.7 GDSA Calculation Archive

In FY 2021, the GDSA team established the GDSA Calculation Archive, a centralized digital repository of GDSA calculations. This archive, hosted on SNL’s internal gitlab site, targets all GDSA milestone calculations and their supporting input files, codes, and workflow. The home page is shown in Figure 3-34.

The main purpose of the archive is for improved internal communication and knowledge capture. The primary goals are:

- To document and archive GDSA calculations in a secure, centralized location
- To allow autonomous sharing of these calculations among GDSA team members
- To provide R&A technical reviewers easy access to supporting materials and calculations
- To provide a springboard for new calculations

Examples of materials in the archive include:

- PFLOTRAN input decks (with paths to data files appropriately updated)
- Data files imported by the PFLOTRAN input decks
- Dakota input files
- Meshes
- Simple PFLOTRAN outputs (e.g., obs files, agg files, mas files) (This specifically excludes certain PFLOTRAN outputs like full .h5 output.)

- Spreadsheets (ideally converted to CSV format unless plots or other non-CSV functionality is used in the spreadsheet)
- README files with appropriate metadata

Metadata in the README files include:

- Originator(s)/author(s)
- PFLOTRAN commit hash (if PFLOTRAN was used)
- Software used and version (if not PFLOTRAN)
- Citation of the deliverable in which materials were used
- Section of the deliverable in which the materials were used
- Description of all files included in the directory
- Possible links to additional materials
- Comments as needed to convey important/useful information

The GDSA Calculation Archive is not a working directory. Updates to preserve the working condition of the materials, however, are welcome. For example, if a PFLOTRAN input deck from a past deliverable becomes incompatible with future code releases, a user could push an updated input deck that produces the same results on an updated version of the code.

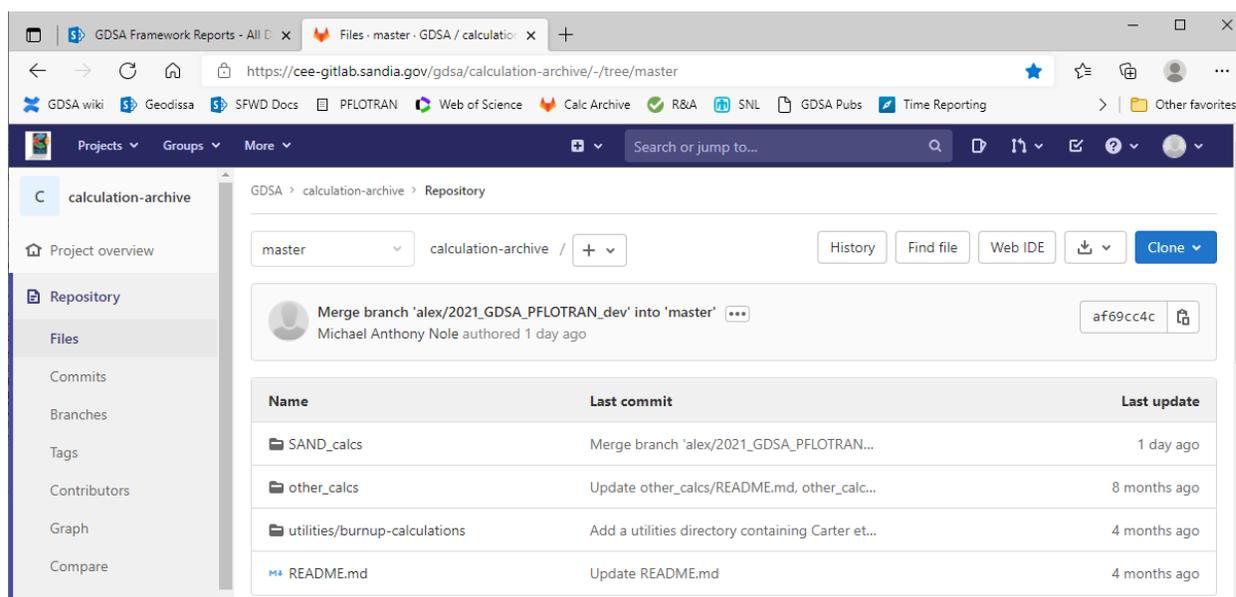


Figure 3-34 GDSA Calculation Archive gitlab home page

3.3 User Group

GDSA Framework is being developed for DOE and its subcontractors. Many of the software components of GDSA Framework are open source, including PFLOTRAN, Dakota, and dfnWorks. These codes are utilized by a community of users from around the world for work related to, and unrelated to, repository performance assessment (e.g., Hammond and Lichtner 2010; Chen et al. 2013; de Vries et al. 2013; Karra

et al. 2014; Gardner et al. 2015; Kumar et al. 2016; Zachara et al. 2016; Avasarala et al. 2017; Trincherro et al. 2017; Shuai et al. 2019).

This section reviews the open source strategy, the PFLOTRAN user community, and how investment in GDSA Framework development is benefitting not only nuclear waste repository performance assessment, but the entire field of subsurface simulation.

3.3.1 Open Source

Open source software licensing governs the free distribution of source code and/or binaries among a group of software developers and users. PFLOTRAN utilizes the GNU LGPL (lesser general public license) which states that the code may be distributed and modified as desired, but any changes to the original source code must be free and publicly available. On the other hand, LGPL allows anyone to link a proprietary third-party library to the code or develop a graphical user interface on top of the code for profit. Further details are provided in Mariner et al. (2019, Section 2.3.4.1).

There are many benefits to open source collaboration, especially when taxpayer funds support much of the code development. First, it encourages collaboration among a diverse team of developers. This collaboration pushes the code to the masses who can help test and debug the code while providing feedback regarding user interaction. Open source provides transparency that exposes implementation details that are often critical for scientific reproducibility and quality assurance. These details are often deliberately or unintentionally omitted from user documentation, journal publications and reports. From a financial standpoint, open source allows developers to pool funds across a diverse set of projects funded in academia, government laboratories or the private sector. In addition, funding that would be spent on licensing fees can be redirected towards development. Finally, although the most fit codes can survive under any licensing option, open source may provide a more level playing field for natural selection to run its course.

PFLOTRAN development is currently supported by multiple developer groups from around the world. DOE is perhaps the largest proponent of PFLOTRAN development through its national laboratories funded by the DOE Offices of Environmental Management, Nuclear Energy, and Science. In addition, private sector companies such as OpenGoSim (opengosim.com) have invested development in support of oil and gas and carbon sequestration efforts, while Amphos21 has developed PFLOTRAN capability for nuclear waste disposal (e.g., de Vries et al. 2013; Iraola et al. 2019).

3.3.2 International User Community

PFLOTRAN's open source licensing and accessible distribution facilitate collaboration amongst a broader U.S. and international community. This broad user community enhances the development of PFLOTRAN by sharing conceptual models, incorporating novel physicochemical algorithms, optimizing code performance, debugging problematic issues, and generating grass-roots publicity, all of which benefit DOE in return.

The PFLOTRAN website at www.pflotran.org directs interested parties to the online documentation and the Bitbucket repository (including source code and documentation build status and code coverage). Developer and user mailing lists are managed through Google Groups.

Estimating the size and extent of the PFLOTRAN user community is relatively difficult due to the inability to track downloads on Bitbucket. However, through Google Analytics, the hits on the PFLOTRAN website are tracked which provides a qualitative estimate (Figure 3-35) and demonstrates that the PFLOTRAN user base is multi-national.

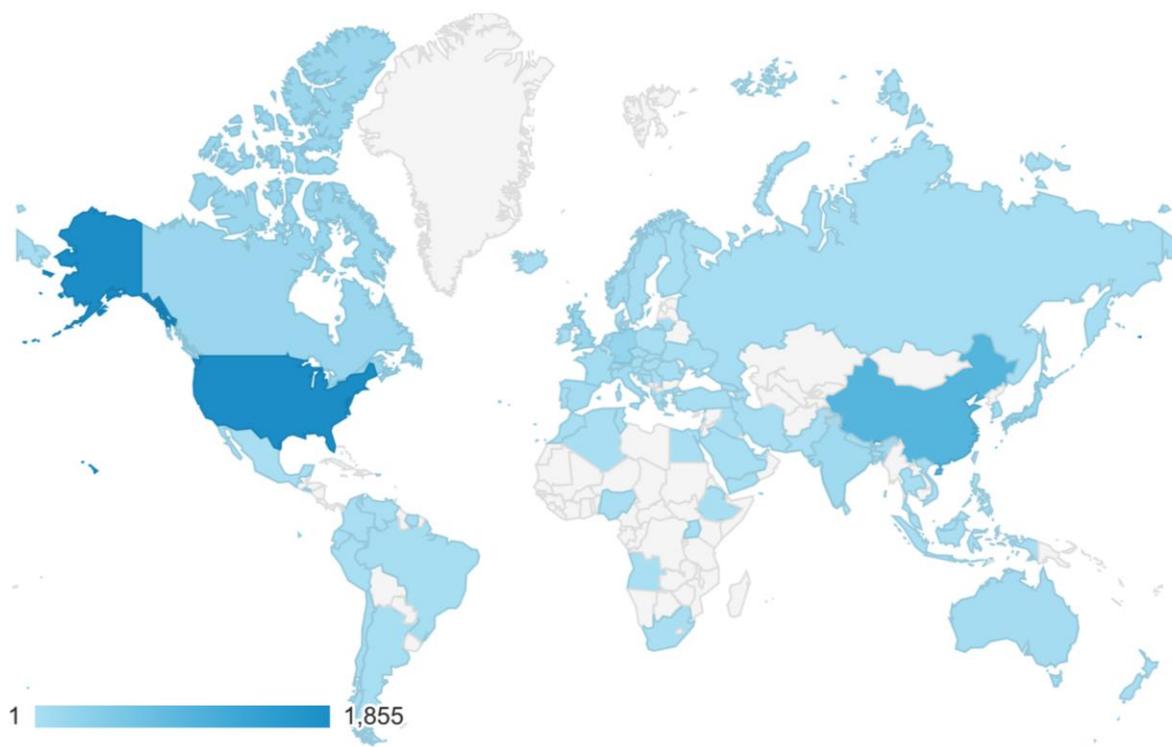


Figure 3-35 User count on the PFLOTRAN website around the world between June 12, 2020 and June 12, 2021, colored by country

3.3.3 Short Courses

This year, a virtual PFLOTRAN short course was conducted over an eight-week period, two hours per week, beginning June 7, 2021. The course had 32 participants across four national labs and six countries. PFLOTRAN developers at Sandia National Laboratories and Pacific Northwest National Laboratory hosted the short course and covered the basics of flow and transport simulations as well as advanced concepts relating to multicontinuum modeling of radionuclide transport and subsurface radioactive waste repository simulations.

3.3.4 GDSA Framework Website

The GDSA team continues to support and develop the GDSA Framework website at <http://pa.sandia.gov/>. This website is publicly available. The home page is shown in Figure 3-36.

The purpose of the website is to:

- Describe GDSA Framework, its capabilities, and the objectives behind its development
- Provide related reports for downloading
- Provide links to software used in GDSA Framework (e.g., PFLOTRAN, Dakota, dfnWorks)
- Identify collaborators involved in GDSA Framework development
- Announce upcoming events (e.g., PFLOTRAN short courses)

- Provide contact information

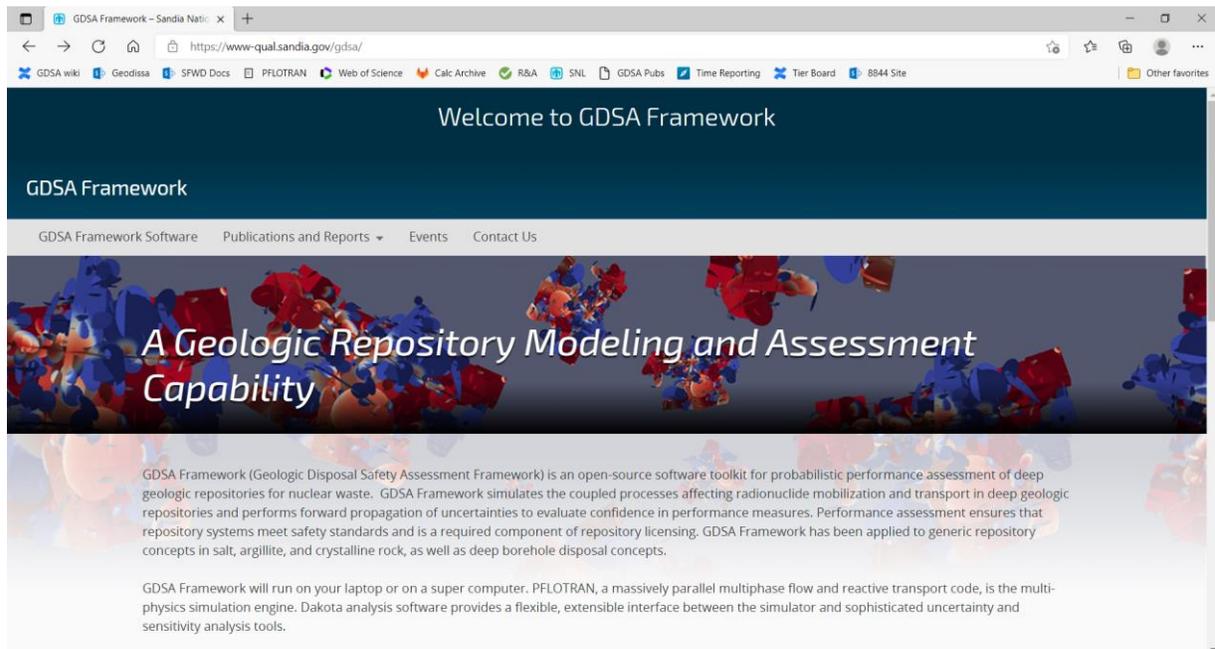


Figure 3-36 GDSA Framework website (<http://pa.sandia.gov/>)

3.3.5 SFWST Document Archive

The GDSA team continues to improve and maintain the SFWST Document Archive (SDA). The SDA is a restricted-access SharePoint website that serves as a document repository for reports generated in:

- NE 81, Office of SFWST
 - Disposal Research (DR)
 - Storage and Transportation (S&T)
- NE 82, Office of Integrated Waste Management (IWM)
- SFWD Annual Meetings

The SDA contains copies of deliverable documents, both unclassified unlimited release and unclassified limited release (ULR), and information about their pedigree (e.g., “downloaded from OSTI,” “best available draft from author,” etc.) and release status (e.g., “ULR,” “internal use only, do not cite or release,” etc.). The archive is searchable and sortable for SFWD participants; however, “Non-Public” archives are restricted as needed. The “Public” archives are available to all SFWD staff, DOE employees, and contractors.

4. DISPOSAL RESEARCH R&D ROADMAP

The 2019 Disposal Research R&D Roadmap Update highlighted progress, priorities, and remaining gaps in disposal research R&D activities (Sevougian et al. 2019). Section 4.1 summarizes ongoing (and gap) GDSA PA activities in the Roadmap Database, and Section 4.2 identifies the FY 2021 GDSA efforts related to those PA activities.

4.1 Findings of the 2019 Roadmap Update for GDSA

Activities defined and tracked in the Roadmap Database are a collection of specific disposal research objectives focused on improving our knowledge of FEPs and how they affect repository performance. These activities include:

- Collecting and measuring the properties of features (e.g., radionuclides, WFs, WP, buffer, DRZ, repository layout, host rock, etc.) and their associated uncertainties
- Identifying and modeling important processes (e.g., flow of heat and groundwater, WP degradation, WF degradation, radionuclide adsorption, buffer evolution, etc.) at small scale and/or in repository simulations
- Estimating the magnitudes, consequences, and probabilities of events that might affect repository performance (e.g., criticality, disruptive events)
- Developing tools and processes to propagate uncertainties in repository performance calculations and to enhance sensitivity analysis

A total of 17 activities were defined for GDSA PA in the 2019 Roadmap update exercise (Sevougian et al. 2019). They are listed in Table 4-1.

Table 4-1 GDSA PA activities in the Roadmap Database of the 2019 Disposal R&D Roadmap update

| Activity | *Gap | Name |
|----------|------|---|
| P-01 | | <i>CSNF repository argillite reference case</i> |
| P-02 | | <i>CSNF repository crystalline reference case</i> |
| P-03 | | <i>CSNF repository bedded salt reference case</i> |
| P-04 | | <i>CSNF repository unsaturated zone (alluvium) reference case</i> |
| P-05 | | <i>Disruptive events</i> |
| P-06 | | <i>(Pseudo) colloid-facilitated transport model</i> |
| P-07 | | <i>Intrinsic colloid model</i> |
| P-08 | * | <i>Other missing FEPs (processes) in PA-GDSA</i> |
| P-09 | | <i>Surface processes and features</i> |
| P-10 | | <i>Uncertainty and sensitivity analysis</i> |
| P-11 | * | <i>Pitzer model</i> |
| P-12 | | <i>WP degradation model framework</i> |
| P-13 | * | <i>Full representation of chemical processes in PA</i> |
| P-14 | | <i>Generic capability development for PFLOTRAN</i> |
| P-15 | * | <i>Species and element properties</i> |
| P-16 | * | <i>Solid solution model</i> |
| P-17 | * | <i>Multi-component gas transport</i> |

4.2 Alignment of FY 2021 GDSA Efforts with 2019 Roadmap Activities

As indicated in Section 3, there were many advances in GDSA PA capabilities in FY 2021. In addition to those advances, there was planning for the near term and longer term, building of infrastructure, and integration across SFWST Disposal Research R&D. This section examines how GDSA efforts in FY 2021 align with the GDSA PA activities identified in the 2019 Roadmap update and how several of the FY 2021 GDSA efforts go beyond the Roadmap activities to support development of GDSA Framework.

The PA activities in the Disposal R&D Roadmap Database (Table 4-1) provide objectives for improving the simulation of physical/chemical/radiological processes in GDSA Framework. Table 4-2 lists the major GDSA efforts in FY 2021 and maps most of them to the PA activities in the Roadmap Database. This mapping indicates that a large number of the PA activities in the Roadmap Database received attention in multiple FY 2021 GDSA efforts. Those that received little or no attention in FY 2021 were colloids (P-06, P-07), Pitzer modeling (P-11), species and elements properties (P-15), and solid solution modeling (P-16).

Not all FY 2021 GDSA efforts listed in Table 4-2 were aimed at addressing GDSA PA activities in the Roadmap Database. Several of them advanced GDSA Framework in other ways. They include advances in workflow processes, code development infrastructure, HPC resources, and planning for future GDSA

Framework development. Those advances directly facilitate the ability of the GDSA development team to address the various GDSA PA objectives.

In addition to listing and mapping the various FY 2021 GDSA efforts to specific GDSA PA activities in the Roadmap Database, Table 4-2 attempts to identify the major parties that participated in the FY 2021 GDSA efforts. GDSA activities commonly require careful integration between SNL GDSA work packages and other parties. Table 4-2 also maps the FY 2021 GDSA efforts to the 5-Year Plan near-term GDSA thrusts (Sassani et al. 2020). The planned GDSA thrusts are discussed in Section 5.

Table 4-2 FY 2021 GDSA activities mapped to GDSA Roadmap activities and 5-Year Plan near-term GDSA thrusts

| FY 2021 GDSA Activities | FY 2021 Participants | Roadmap Database PA Activity | Near-Term GDSA Thrusts |
|--|--------------------------------|-------------------------------------|-------------------------------|
| 5-Year disposal R&D plan update (Section 5.2) | SNL | - | All |
| Agile/Jira system operation (Section 3.1.1.1) | SNL | - | G03 |
| Biosphere modeling (Section 3.1.3.8) | PNNL, SNL | P-08, P-09 | G01, G04 |
| Buffer and DRZ evolution (Sections 3.1.3.2 and 3.1.3.5) | SNL | P-01, P-14 | G01 |
| DECOVALEX-2023 Task F: Performance assessment (Section 3.2.4) | SNL, LBNL, LANL, International | P-02, P-03, P-05, P-12 | G02, G04 |
| dfnWorks integration (Sections 3.2.3.1, 3.2.3.2, 3.2.4.1, and 3.2.6.2) | SNL, LANL | P-02, P-10, P-14 | G01, G02 |
| Direct DPC disposal modeling (Section 3.1.3.7) | SNL, ORNL | P-03, P-04, P-14 | G01, G04 |
| FEPs in salt repositories (Freeze et al. 2020) | SNL, LBNL | P-03, P-08 | G04 |
| Fuel matrix degradation (Section 3.1.3.3) | SNL, ANL | P-13, P-14 | G01 |
| GDSA Calculation Archive (Section 3.2.7) | SNL | - | G03 |
| Geologic Framework Model (Section 3.2.5) | SNL, LANL, INL | P-01, P-02, P-04, P-09 | G03, G04, G05 |
| Graphical workflow interface for GDSA Framework simulations (Section 3.2.1) | SNL | - | G02, G03 |
| HPC resources utilization (Section 3.1.1.3) | SNL | - | G03 |
| Multi-continuum transport development (Section 3.1.3.4) | SNL, LANL | P-02, P-14 | G01 |
| Multifidelity uncertainty methods (Section 3.2.3.1) | SNL | P-02, P-10 | G02 |
| PFLOTRAN convergence (Section 3.1.2) | SNL | P-01, P-02, P-03, P-04, P-17 | G01 |
| Process model coupling requirements, guidelines (Section 3.1.3.1) | SNL | - | G03 |
| QA toolbox and test suite development (Section 3.1.1.2) | SNL | - | G03 |
| Repository reference case development (crystalline, argillite, salt, unsaturated alluvium) (Section 3.2.6) | SNL | P-01, P-02, P-03, P-04, P-10 | G02, G04 |
| Re-saturation geomechanics modeling (Section 3.2.6.1) | SNL, LBNL | P-14 | G01, G04 |
| Smectite-to-illite transition (Section 3.1.3.6) | SNL | P-14 | G01 |
| Thermal conductivity enhancements (Section 3.1.3.5) | SNL | P-14 | G01, G04 |
| Uncertainty and sensitivity analysis (UQ/SA) (Section 3.2.3) | SNL | P-02, P-10 | G02, G04 |
| UQ/SA international (Section 3.2.3) | SNL, International | P-01, P-02, P-10 | G02, G04 |
| Voronoi meshing and simulation (Section 3.2.2) | SNL | P-01, P-02, P-03, P-04, P-09 | G01, G04 |
| Waste package degradation (Section 3.1.3.2) | SNL | P-12, P-13 | G01 |

5. FIVE-YEAR PLAN FOR DISPOSAL RESEARCH R&D

5.1 5-Year Plan

In FY 2020, DOE requested development of a 5-year plan for high priority activities of the SFWST Campaign Disposal Research (DR) R&D activities. As requested, that plan was prepared and delivered to DOE in July of 2020 (Sassani et al. 2020). In the plan, current and proposed activities for each DR R&D technical area were evaluated and categorized in terms of near-term and longer-term thrusts. In FY 2021, that plan, as discussed in Section 5.2, was updated.

The original FY 2020 5-year plan identified five primary thrusts for the near term (1-2 years) for GDSA:

- Advanced coupled process simulation capabilities (G01),
- State-of-the-art uncertainty and sensitivity analysis (G02),
- Traceable, user-friendly workflow for GDSA Framework (G03),
- Repository systems analysis for various disposal concepts and selected host rocks (G04), and
- Development of geologic models with interactive, web-based visualization (G05).

Longer-term (3-5 years) thrusts in the plan included multi-fidelity modeling, in-package chemistry, gas flow in the EBS, cement seal evolution, new repository designs, and preparation for site applications. An additional GDSA focus area identified in Section 3 of that plan addressed in-drift coupled chemistry modeling and major chemical reactions with materials over appropriate temperature ranges.

5.2 5-Year Plan Update

The updated 5-year plan is documented in Sassani et al. (2021). It follows the original plan in structure and general focus but is updated with perceptible changes due to accomplishments in FY 2021 and evolving areas of interest.

The five primary thrusts for the near term (1-2 years) for GDSA (i.e., identified as G01 through G05 above) are generally unchanged. Their updated descriptions are provided verbatim in Appendix C. The changes include:

- *G01 – Advanced simulation capability.* The LANL software *dfnWorks* was explicitly added as a capability to include in the development of GDSA Framework. Development of advanced solvers and physics-based convergence criteria was removed.
- *G02 – Uncertainty and sensitivity analysis.* Evaluation of model form uncertainty was added. Evaluation of the reliability of methods dependent on surrogate models through techniques such as cross-validation and development of quantitative metrics for assessing goodness of surrogates was also added. Work toward international consensus on best practices was replaced with continuation of a leadership role in international collaborations on the subject.
- *G03 – Workflow.* Integration of GDSA Framework with the Online Waste Library OWL was added.
- *G04 – Repository systems analysis.* A near term priority for the development of salt and crystalline reference cases was added due to involvement in Task F of DECOVALEX-2023. Mention of the US/German joint FEPs project for salt was removed due to completion in FY 2021.

- *G05 – Geologic modeling.* No changes.

The longer-term GDSA thrusts were unchanged from the original plan. Section 5.3 addresses how FY 2021 GDSA activities and Roadmap activities align with the near-term thrusts.

5.3 Alignment of the Near-Term GDSA Thrusts with FY 2021 GDSA Activities and GDSA Roadmap Activities

To evaluate the coverage of the near-term GDSA thrusts by FY 2021 GDSA activities, the thrusts were linked to FY 2021 GDSA activities in Table 4-1. The thrust identifier (G01 through G05) was linked to FY 2021 GDSA activities in the last column of Table 4-2 if the activity aligned with the thrust.

As indicated in Table 4-2 each thrust was addressed to some degree in FY 2021. To link the FY 2021 GDSA infrastructure efforts to a thrust, a broader interpretation of the G03 Workflow thrust was used. Certainly, it should be noted that while each of the thrusts garnered some attention in FY 2021, there was no attempt to quantify how much of the scope of each thrust was addressed in FY 2021.

To evaluate the linkages between the GDSA thrusts of the 5-Year Plan and GDSA Roadmap activities, each activity was evaluated for its ability to address a GDSA thrust. As shown in Appendix D, each near-term thrust was found to have at least one GDSA Roadmap activity linked to it, and each GDSA Roadmap activity was linked to at least one GDSA thrust. Several GDSA thrusts, especially those involving the development of new modeling capability, were linked to a large number of GDSA Roadmap activities.

At this stage of the 5-Year Plan, it is unclear how much the mapping of efforts, Roadmap activities, and thrusts (e.g., Appendix D) can help guide planning and production efforts in the years ahead. However, this type of mapping does provide a fairly organized way to better understand how each thrust and Roadmap activity is being addressed.

6. CONCLUSIONS

This report describes FY 2021 advances of the Geologic Disposal Safety Assessment (GDSA) performance assessment (PA) development groups of the SFWST Campaign. The common mission of these groups is to develop a geologic disposal system modeling capability for nuclear waste that can be used to assess probabilistically the performance of generic disposal options and generic sites. The developing capability, called GDSA Framework, employs high-performance computing capable codes PFLOTRAN and Dakota.

The advances in PA modeling capabilities allow for emulation of fuel matrix degradation processes at each waste package in a repository simulation, simulation of changes to buffer and disturbed rock zone properties over time, improved simulation of heat flow due to new temperature dependencies included in the calculation of thermal conductivity, improved PFLOTRAN convergence for multiphase systems and dry out, and more detailed assessment of total system performance. New methods allow for quantitative bulk characterization of flow and transport in regions affected by discrete fracture networks and for simulating the interaction of fracture and matrix processes. New modeling approaches include advanced meshing capabilities, a new plan for simulating buffer erosion and waste package corrosion, and development of a comprehensive biosphere model. Progress in simulation workflow, quality assurance (QA) workflow, process model coupling workflow, and other forms of supporting infrastructure is expected to greatly facilitate future model development, user-friendliness, and user adoption.

An important responsibility of the GDSA team is to integrate with disposal R&D activities across the SFWST Campaign to ensure that R&D activities support the parts of the generic safety cases being developed. In FY 2021, the GDSA team continued to participate with other scientists and engineers at LANL, LBNL, PNNL, ORNL, INL, ANL, DOE, and SNL in the development of new high-temperature model capabilities needed for possible direct disposal of dual-purpose canisters, dfnWorks, Geologic Framework Models, FMD model integration, DECOVALEX-2023 Task F performance assessment, and advanced biosphere modeling.

Each year, GDSA Framework improves as additional modelers and programmers from around the world use, apply, and contribute to its development. GDSA Framework can be shared because the primary codes, PFLOTRAN and Dakota, are open source, available for free download, and have supporting documentation online. Outreach and collaborations support a primary objective of the GDSA work package by facilitating testing of, and feedback on, PFLOTRAN and GDSA Framework and by increasing the likelihood outside users will contribute directly to code development in the future.

The ability to simulate increasingly complex repository reference cases continues to affirm that HPC-capable codes can be used to simulate important multi-physics couplings directly in a total system safety assessment demonstration. Reference-case-repository applications show that PFLOTRAN and its coupled codes can simulate complex coupled processes in a multi-kilometer domain while simultaneously simulating sub-meter-scale coupled behavior in the vicinity of each modeled waste package. Continued development will further enhance the preparedness of GDSA Framework for application in the future when transitioning to a program with potential sites.

7. REFERENCES

- Adams, B.M., K.R. Dalbey, M.S. Eldred, L.P. Swiler, W.J. Bohnhoff, J.P. Eddy, D.M. Vigil, P.D. Hough and S. Lefantzi (2012). DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 5.2+ User's Manual. Sandia National Laboratories, Albuquerque, New Mexico.
- Adams, B.M., M.S. Ebeida, M.S. Eldred, J.D. Jakeman, L.P. Swiler, W.J. Bohnhoff, K.R. Dalbey, J.P. Eddy, K.T. Hu, D.M. Vigil, L.E. Baumann and P.D. Hough (2013). Dakota, a Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis, Version 5.3.1+ Theory Manual. Sandia National Laboratories, Albuquerque, New Mexico.
- Avasarala, S., P.C. Lichtner, A. S. Ali, R. Gonzalez-Pinzon, J.M. Blake and J.M. Cerrato (2017). "Reactive Transport of U and V from Abandoned Uranium Mine Wastes." *Environmental Science & Technology* 51(21): 12385-12393.
- Balay, S., J. Brown, K. Buschelman, V. Eijkhout, W.D. Gropp, D. Kaushik, M.G. Knepley, L. Curfman McInnes, B.F. Smith and H. Zhang (2013). PETSc Users Manual. Argonne, Illinois, Argonne National Laboratory.
- Berglund, S., Bosson, E., Selroos, JO, and Sassner, M. (2013). Identification and Characterization of Potential Discharge Areas for Radionuclide Transport by Groundwater from a Nuclear Waste Repository in Sweden. *AMBIO*. 42:435–446.
- Brady, P.V., B.W. Arnold, G.A. Freeze, P.N. Swift, S.J. Bauer, J.L. Kanney, R.P. Rechard, and J.S. Stein (2009). Deep borehole disposal of high-level radioactive waste. SAND2009-4401. Sandia National Laboratories, Albuquerque, New Mexico.
- Chen, X., G. Hammond, C. Murray, M. Rockhold, V. Vermeul and J. Zachara (2013). "Applications of ensemble-based data assimilation techniques for aquifer characterization using tracer data at Hanford 300 area." *Water Resources Research* 49: 7064-7076.
- Chen, X., H. Murakami, M. Hahn, G. E. Hammond, M.L. Rockhold, J.M. Zachara and Y. Rubin (2012). "Three-Dimensional Bayesian Geostatistical Aquifer Characterization at the Hanford 300 Area using Tracer Test Data." *Water Resources Research* 48.
- Condon, C.A., B. A. Napier, S. Ghosh, W.C. Weaver and C.B. Varnum-Lowry (2020). GDSA Biosphere Model Software Requirements Document. M3SF-20PN010304091. PNNL-30280. Richland, Washington, Pacific Northwest National Laboratory.
- Dale, T.N. (1923). The commercial granites of New England. U.S. Geological Survey. Bulletin 738. 488 p.
- de Vries, L. M., J. Molinero, H. Ebrahimi, U. Svensson, P. Lichtner and E. Abarca (2013). Regional Scale HPC Reactive Transport Simulation of Nuclear Spent Fuel Repository in Forsmark, Sweden. Joint 8th International Conference on Supercomputing in Nuclear Applications (SNA) / 4th Monte Carlo Meeting (MC), Paris, France.
- DOE (2011). Used Fuel Disposition Campaign Disposal Research and Development Roadmap. Washington, DC, Fuel Cycle Technologies, Office of Nuclear Energy, US Department of Energy.
- DOE (2013). Strategy for the Management and Disposal of Used Nuclear Fuel and High-Level Radioactive Waste, January 2013. Washington, DC. <http://energy.gov/downloads/strategy-management-and-disposal-used-nuclear-fuel-and-high-level-radioactive-waste>, U.S. Department of Energy.
- Follin S., Hartley L., Jackson P., Roberts D., Marsic N. (2008). Conceptual model development and numerical modelling using CONNECTFLOW, Forsmark modelling stage 2.3. SKB R-08-23. Svensk Kärnbränslehantering AB.
- Freeze, G., M. Voegele, P. Vaughn, J. Prouty, W.M. Nutt, E. Hardin and S.D. Sevougian (2013b). Generic Deep Geologic Disposal Safety Case. Sandia National Laboratories, Albuquerque, New Mexico.

- Freeze, G., P.E. Mariner, J.A. Blink, F.A. Caporuscio, J.E. Houseworth and J.C. Cunnane (2011). Disposal System Features, Events, and Processes (FEPs): FY11 Progress Report. Sandia National Laboratories, Albuquerque, New Mexico.
- Freeze, G., Stein, E.R., Brady, P.V., Lopez, C., Sassani, D. (2019). Deep Borehole Disposal Safety Case SAND2019-1915. Sandia National Laboratories, Albuquerque, New Mexico.
- Freeze, G., Stein, E.R., Price, L., MacKinnon, R., and Tillman, J. (2016). Deep Borehole Disposal Safety Analysis. FCRD-UFD-2016-000075, SAND2016-10949R. Sandia National Laboratories, Albuquerque, New Mexico.
- Freeze, G., W.P. Gardner, P. Vaughn, S.D. Sevougian, P. Mariner, V. Mousseau and G. Hammond (2013a). Enhancements to the Generic Disposal System Modeling Capabilities. Sandia National Laboratories, Albuquerque, New Mexico.
- Gardner, W.P., G. Hammond and P. Lichtner (2015). "High Performance Simulation of Environmental Tracers in Heterogeneous Domains." *Groundwater* 53: 71-80.
- Hammond, G.E. and P. Lichtner (2010). "Field-scale modeling for the natural attenuation of uranium at the Hanford 300 area using high performance computing." *Water Resources Research* 46.
- Hammond, G.E., P.C. Lichtner and M.L. Rockhold (2011b). "Stochastic simulation of uranium migration at the Hanford 300 Area." *Journal of Contaminant Hydrology* 120-121: 115-128.
- Hammond, G.E., P.C. Lichtner, C. Lu and R.T. Mills (2011a). PFLOTRAN: Reactive Flow and Transport Code for Use on Laptops to Leadership-Class Supercomputers. *Groundwater Reactive Transport Models*. F. Zhang, G. T. Yeh and J. Parker, Bentham Science Publishers.
- Hammond, G.E., P.C. Lichtner, R.T. Mills and C. Lu (2008). "Toward petascale computing in geosciences: application to the Hanford 300 Area." *Journal of Physics Conference Series* 125: 12051-12051.
- Hammond, G., P. Lichtner and C. Lu (2007). "Subsurface multiphase flow and multicomponent reactive transport modeling using high performance computing." *Journal of Physics: Conference Series* 78: 1-10.
- Hansen, B.P., Stone, J.R., and Lane, J.W. (1999). Characteristics of fractures in crystalline bedrock determined by surface and borehole geophysical surveys, eastern surplus superfund site, Meddybemps, Maine. *Water-Resources Investigations Report*. 99-4050. 27 p.
- Huang, W.L., Longo, J.M., & Pevear, D.R. (1993). An experimentally derived kinetic model for smectite-to-illite conversion and its use as a geothermometer. *Clays and Clay Minerals*, 41(2), 162-177.
- IAEA (2003). Reference Biospheres for Solid Radioactive Waste Disposal. Vienna, Austria, International Atomic Energy Agency.
- IAEA (2006). Geological Disposal of Radioactive Waste. Safety Requirements No. WS-R-4. Vienna, Austria, International Atomic Energy Agency.
- Iraola, A., P. Trinchero, S. Karra and J. Molinero (2019). Assessing dual continuum method for multicomponent reactive transport. *Computers & Geosciences* 130: 11-19.
- Jerden, J., G. Hammond, J. M. Copple, T. Cruse and W. Ebert (2015b). Fuel Matrix Degradation Model: Integration with Performance Assessment and Canister Corrosion Model Development. Washington, DC, US Department of Energy.
- Jerden, J., J.M. Copple, K. E. Frey and W. Ebert (2015a). Mixed Potential Model for Used Fuel Dissolution - Fortran Code. O. o. U. N. F. Disposition. Washington, DC, US Department of Energy.
- Jerden, J., V.K. Gattu and W. Ebert (2018). Update on Validation and Incorporation of a New Steel Corrosion Module into Fuel Matrix Degradation Model. Illinois, Argonne National Laboratory.
- Johansson P-O. 2008. Description of surface hydrology and near-surface hydrogeology at Forsmark. Site descriptive modelling, SDM-Site Forsmark. SKB R-08-08. Svensk Kärnbränslehantering AB.
- Jove Colon, C.F., P. Weck, L. Zheng, J. Rutqvist, C.I. Steefel, K. Kim, S. Nakagawa, J. Houseworth, J. Birkholzer, F.A. Caporuscio, M. Cheshire, M.S. Rearick, M.K. McCarney, M. Zavarin, A.B. Benedicto, M.S. Kersting, J.L. Jerden, K.E. Frey, J.M. Copple, and W.L. Ebert (2014).

- Evaluation of Used Fuel Disposition in Clay-Bearing Rock. FCRD-UFD-2014-000056. SAND2014-18303R. Sandia National Laboratories, Albuquerque, New Mexico.
- Karra, S., S.L. Painter and P.C. Lichtner (2014). "Three-phase numerical model for subsurface hydrology in permafrost-affected regions (PFLOTTRAN-ICE v1.0)." *Cryosphere* 8(5): 1935-1950.
- Kumar, J., N. Collier, G. Bisht, R. T. Mills, P. E. Thornton, C. M. Iversen and V. Romanovsky (2016). "Modeling the spatiotemporal variability in subsurface thermal regimes across a low-relief polygonal tundra landscape." *Cryosphere* 10(5): 2241-2274.
- LaForce, T., Chang, K.W., Perry, F.V., Lowry, T.S., Basurto, E., Jayne, R., Brooks, D., Jordan, S., Stein, E.R., Leone, R., Nole, M. (2020) GDSA Repository Systems Analysis Investigations in FY 2020. SAND2020-12028 R. Sandia National Laboratories, Albuquerque, New Mexico.
- LaForce, T., E. Basurto, K.W. Chang, R. Jayne, R. Leone, M. Nole, F.V. Perry, and E. Stein (2021, in progress). GDSA Repository Systems Analysis Investigations in FY2021. Sandia National Laboratories, Albuquerque, New Mexico.
- Lichtner, P.C. and G.E. Hammond (2012). Quick Reference Guide: PFLOTTRAN 2.0 (LA-CC-09-047) Multiphase-Multicomponent-Multiscale Massively Parallel Reactive Transport Code. Los Alamos, New Mexico, Los Alamos National Laboratory.
- Lichtner, P.C. (2000). Critique of Dual Continuum Formulations of Multicomponent Reactive Transport in Fractured Porous Media, Ed. Boris Faybishenko, *Dynamics of Fluids in Fractured Rock*, Geophysical Monograph 122, 281–298.
- Lu, C. and P. C. Lichtner (2007). "High resolution numerical investigation on the effect of convective instability on long term CO₂ storage in saline aquifers." *Journal of Physics Conference Series* 78: U320-U325.
- Mariner, P. E., E. R. Stein, J. M. Frederick, S. D. Sevougian and G. E. Hammond (2017). Advances in Geologic Disposal System Modeling and Shale Reference Cases. SFWD-SFWST-2017-000044, SAND2017-10304 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Mariner, P. E., E. R. Stein, J. M. Frederick, S. D. Sevougian, G. E. Hammond and D. G. Fascitelli (2016). Advances in Geologic Disposal System Modeling and Application to Crystalline Rock. FCRD-UFD-2016-000440, SAND2016-9610 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Mariner, P. E., E. R. Stein, S. D. Sevougian, L. J. Cunningham, J. M. Frederick, G. E. Hammond, T. S. Lowry, S. Jordan and E. Basurto (2018). Advances in Geologic Disposal Safety Assessment and an Unsaturated Alluvium Reference Case. SFWD-SFWST-2018-000509, SAND2018-11858 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Mariner, P. E., L. A. Connolly, L. J. Cunningham, B. J. Debusschere, D. C. Dobson, J. M. Frederick, G. E. Hammond, S. H. Jordan, T. C. LaForce, M. A. Nole, H. D. Park, F. V. Perry, R. D. Rogers, D. T. Seidl, S. D. Sevougian, E. R. Stein, P. N. Swift, L. P. Swiler, J. Vo and M. G. Wallace (2019). Progress in Deep Geologic Disposal Safety Assessment in the U.S. since 2010. SAND2019-12001 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Mariner, P. E., T. M. Berg, K. W. Chang, B. J. Debusschere, R. C. Leone and D. T. Seidl (2020b). Surrogate Model Development of Spent Fuel Degradation for Repository Performance Assessment. SAND2020-10797 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Mariner, P. E., W. P. Gardner, G. E. Hammond, S. D. Sevougian and E. R. Stein (2015). Application of Generic Disposal System Models. FCRD-UFD-2015-000126, SAND2015-10037R. Sandia National Laboratories, Albuquerque, New Mexico.
- Mariner, P.E., M.A. Nole, E. Basurto, T.M. Berg, K.W. Chang, B.J. Debusschere, A.C. Eckert, M.S. Ebeida, M. Gross, G.E. Hammond, J. Harvey, S.H. Jordan, K.L. Kuhlman, T.C. LaForce, R.C. Leone, W.C. McLendon III, M.M. Mills, H.D. Park, F.V. Perry, A. Salazar III, D.T. Seidl, S.D. Sevougian, E.R. Stein, and L.P. Swiler (2020a). Advances in GDSA Framework Development and Process Model Integration. M2SF-20SN010304042. SAND2020-10787 R. Sandia National Laboratories, Albuquerque, New Mexico.

- Meacham, P. G., D. R. Anderson, E. J. Bonano and M. G. Marietta (2011). Sandia National Laboratories Performance Assessment Methodology for Long-Term Environmental Programs: The History of Nuclear Waste Management. Sandia National Laboratories, Albuquerque, New Mexico.
- MGS (Maine Geological Survey) Open Data (2017). Maine Surficial Geology 500K Units, Available at <https://mgs-maine.opendata.arcgis.com/datasets/maine::maine-surficial-geology-500k-units/about>.
- Mills, R., C. Lu, P. C. Lichtner and G. Hammond (2007). Simulating subsurface flow and transport on ultrascale computers using PFLOTRAN. 3rd Annual Scientific Discovery through Advanced Computing Conference (SciDAC 2007), Boston, Journal of Physics Conference Series.
- Navarre-Sitchler, A., R. M. Maxwell, E. R. Siirila, G. E. Hammond and P. C. Lichtner (2013). Elucidating geochemical response of shallow heterogeneous aquifers to CO₂ leakage using high-performance computing: implications for monitoring CO₂ sequestration. *Advances in Water Resources* 53: 44-55.
- Nole, M., R.C. Leone, H.D. Park, M. Paul, A. Salazar, G.E. Hammond, and P.C. Lichtner (2021). PFLOTRAN Development FY2021. M3SF-21SN010304072. SAND2021-8709 R. Sandia National Laboratories, Albuquerque, New Mexico.
- OECD (2004). Post-Closure Safety Case for Geological Repositories Nature and Purpose. Paris, France, Organisation for Economic Co-Operation and Development, Nuclear Energy Agency.
- Orient, G., R. Clay, E. Friedman-Hill and E. Hoffman (2020). Next Generation Workflow – an Open Source Software System for Computational Modeling to Support Agile VVUQ. ASME V&V Symposium, May 20-22, 2020.
- Posiva (2013). Safety Case for the Disposal of Spent Nuclear Fuel at Olkiluoto - Models and Data for the Repository System 2012. POSIVA 2013-01. Posiva Oy, Olkiluoto, Eurajoki, Finland.
- Price, L.L., A.A. Alsaed, A. Barela, P.V. Brady, F. Gelbard, M.B. Gross, M. Nole, J.L. Prouty, K. Banerjee, S. Bhatt, G.G. Davidson, Z. Fang, R. Howard, S.R. Johnson, S.L. Painter, and M. Swinney (2019b). Preliminary Analysis of Postclosure DPC Criticality Consequences. M2SF-20SN010305061. SAND2020-4106. Sandia National Laboratories, Albuquerque, New Mexico.
- Price, L.L., A.A. Alsaed, P.V. Brady, M.B. Gross, E.L. Hardin, M. Nole, J.L. Prouty, K. Banerjee and G.G. Davidson (2019a). Postclosure Criticality Consequence Analysis—Scoping Phase. M3SF-19SN010305061. Sandia National Laboratories, Albuquerque, New Mexico.
- Rechard, R.P. (1995). Performance Assessment of the Direct Disposal in Unsaturated Tuff of Spent Nuclear Fuel And High-Level Waste Owned by US Department of Energy. Sandia National Laboratories, Albuquerque, New Mexico. 1, 2, and 3.
- Rechard, R.P. (2002). General approach used in the performance assessment for the Waste Isolation Pilot Plant. *Scientific Basis for Nuclear Waste Management XXV*, Boston, Massachusetts, Materials Research Society.
- Rechard, R.P. and C.T. Stockman (2014). "Waste degradation and mobilization in performance assessments of the Yucca Mountain disposal system for spent nuclear fuel and high-level radioactive waste." *Reliability Engineering and System Safety* 122(2): 165-188.
- Rechard, R.P. and M.S. Tierney (2005). "Assignment of probability distributions for parameters in the 1996 performance assessment for the Waste Isolation Pilot Plant, Part 1: Description of process." *Reliability Engineering and System Safety* 88(1): 1-32.
- Ridgway, E.M. (2020). Dakota. GUI Version 6.12 User Manual – Next-Gen Workflow. <https://dakota.sandia.gov/content/next-gen-workflow>. Accessed August 5, 2020.
- Sassani, D., J. Birkholzer, R. Camphouse, G. Freeze and E. Stein (2020). SFWST Disposal Research R&D 5-Year Plan - Draft Report. SAND2020-9053 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Sassani, D., J. Birkholzer, R. Camphouse, G. Freeze and E. Stein (2021). SFWST Disposal Research R&D 5-Year Plan – FY2021 Update. Draft Report. M2SF-21SN010304054. Draft Report - SAND2021-12491 R. Sandia National Laboratories, Albuquerque, New Mexico.

- Sevougian, S. D., E. R. Stein, T. LaForce, F. V. Perry, T. S. Lowry, L. J. Cunningham, M. Nole, C. B. Haukwa, K. W. Chang and P. E. Mariner (2019b). GDSA Repository Systems Analysis Progress Report. SAND2019-5189 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Sevougian, S. D., E. R. Stein, T. LaForce, F. V. Perry, T. S. Lowry, M. Nole and K. W. Chang (2019c). GDSA Repository Systems Analysis FY 2019 Update. Sandia National Laboratories, Albuquerque, New Mexico.
- Sevougian, S. D., G. A. Freeze, P. Vaughn, P. Mariner and W. P. Gardner (2013). Update to the Salt R&D Reference Case. FCRD-UFD-2013-000368. SAND2013-8255P. Sandia National Laboratories, Albuquerque, New Mexico.
- Sevougian, S. D., G. A. Freeze, W. P. Gardner, G. E. Hammond and P. E. Mariner (2014). Performance Assessment Modeling and Sensitivity Analyses of Generic Disposal System Concepts. Sandia National Laboratories, Albuquerque, New Mexico.
- Sevougian, S. D., P. E. Mariner, L. A. Connolly, R. J. MacKinnon, R. D. Rogers, D. C. Dobson and J. L. Prouty (2019a). DOE SFWST Campaign R&D Roadmap Update, Rev. 1. SAND2019-9033R. Sandia National Laboratories, Albuquerque, New Mexico.
- Sevougian, S.D., G.A. Freeze, M.B. Gross, J. Lee, C.D. Leigh, P. Mariner, R.J. MacKinnon, and P. Vaughn (2012). TSPA Model Development and Sensitivity Analysis of Processes Affecting Performance of a Salt Repository for Disposal of Heat-Generating Nuclear Waste. FCRD-UFD-2012-000320 Rev. 0. U.S. Department of Energy, Office of Nuclear Energy, Used Nuclear Fuel Disposition, Washington, D.C.
- Sevougian, S.D., Stein, E.R., Gross, M.B., Hammond, G.E., Frederick, J.M., and Mariner, P.E. (2016). Status of Progress Made Toward Safety Analysis and Technical Site Evaluations for DOE Managed HLW and SNF. FCRD-UFD-2016-000082. SAND2016-11232R. Sandia National Laboratories, Albuquerque, New Mexico.
- Shuai, P., X. Y. Chen, X. H. Song, G. E. Hammond, J. Zachara, P. Royer, H. Y. Ren, W. A. Perkins, M. C. Richmond and M. Y. Huang (2019). "Dam Operations and Subsurface Hydrogeology Control Dynamics of Hydrologic Exchange Flows in a Regulated River Reach." *Water Resources Research* 55(4): 2593-2612.
- Somerton, W. H., Keese, J. A., & Chu, S. L. (1974). Thermal behavior of unconsolidated oil sands. *Society of Petroleum Engineers Journal*, 14(05), 513-521.
- Stein, E.R., C. Bryan, D.C. Dobson, E.L. Hardin, C. Jové-Colón, C.M. Lopez, E.N. Matteo, S. Mohanty, M. Pendleton, F.V. Perry, J.L. Prouty, D.C. Sassani, Y. Wang, J. Rutqvist, L. Zheng, K.B. Sauer, F. Caporuscio, R. Howard, A. Adeniyi, K. Banerjee, and R. Josep (2020). Disposal Concepts for a High-Temperature Repository in Shale. M3SF-21SN010304064. SAND2020-12471 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Sudicky, E.A., and E. Frind (1982) Contaminant transport in fractured porous media: Analytical solution for a system of parallel fractures. *Water Resources Research*, 18 (3), 1634–1642.
- Swiler, L. P., E. Basurto, D. M. Brooks, A. C. Eckert, P. E. Mariner, T. Portone and E. R. Stein (2020). Status Report on Uncertainty Quantification and Sensitivity Analysis Tools in the Geologic Disposal Safety Assessment (GDSA) Framework. SAND2020-10802 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Swiler, L. P., J. C. Helton, E. Basurto, D. M. Brooks, P. E. Mariner, L. M. Moore, S. Mohanty, S. D. Sevougian and E. R. Stein (2019). Status Report on Uncertainty Quantification and Sensitivity Analysis Tools in the Geologic Disposal Safety Assessment (GDSA) Framework. SAND2019-13835 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Swiler, L.P., E. Basurto, D.M. Brooks, A.C. Eckert, R. Leone, P.E. Mariner, T. Portone, M. L. Smith and E.R. Stein (2021). Uncertainty and Sensitivity Analysis Methods and Applications in the GDSA Framework (FY2021). M3SF-21SN010304042. SAND2021-9903R. Sandia National Laboratories, Albuquerque, New Mexico.

- Tang, D.H., E.O. Frind, and E.A. Sudicky (1981). "Contaminant transport in fracture porous media: analytical solution for a single fracture". *Water Resources Research*, 17(3), 555-564. doi: 10.1029/WR017i003p00555
- Thompson, W.B. (1987). The Presumpscot Formation in Southwestern Maine. In *Geologic and Geotechnical Characteristics of the Presumpscot Formation Maine's Glaciomarine "Clay" Symposium*, Andrews, D.W., Thompson, W.D., Sandford, T.C., and Novak, I.D., Eds. Augusta, ME. 20 March 1987. 22 p.
- Thompson, WB. (2015). *Surficial geology handbook for southern Maine: Maine Geological Survey. Bulletin 44.* 97 p. Maine Geological Survey Publications 2. http://digitalmaine.com/mgs_publications/2.
- Trincherro, P., I. Puigdomenech, J. Molinero, H. Ebrahimi, B. Gylling, U. Svensson, D. Bosbach and G. Deissmann (2017). "Continuum-based DFN-consistent numerical framework for the simulation of oxygen infiltration into fractured crystalline rocks." *Journal of Contaminant Hydrology* 200: 60-69.
- Wang, Y., E. N. Matteo, J. Rutqvist, J. A. Davis, L. Zheng, J. Houseworth, J. Birkholzer, T. Dittrich, C. W. Gable, S. Karra, N. Makedonska, S. Chu, D. Harp, S. L. Painter, P. W. Reimus, F. V. Perry, P. Zhao, J. Begg, M. Zavarin, S. J. Tumey, Z. Dai, A. B. Kersting, J. L. Jerden, K. E. Frey, J. M. Copple, and W. L. Ebert (2014). *Used Fuel Disposal in Crystalline Rocks Status and FY14 Progress.* FCRD-UFD-2014-000060. SAND2014-17992 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Zachara, J. M., X. Y. Chen, C. Murray and G. Hammond (2016). "River stage influences on uranium transport in a hydrologically dynamic groundwater-surface water transition zone." *Water Resources Research* 52(3): 1568-1590.
- Zheng, L. C. Jové Colón, M. Bianchi, and J. Birkholzer (2014). *Generic Argillite/Shale Disposal Reference Case.* FCRD-UFDC-2014-000319. LBNL-6709E. Lawrence Berkeley National Laboratory, Berkeley, California.

(This page is intentionally blank.)

Appendix A. FUEL MATRIX DEGRADATION MODELING IN FORTRAN

APPENDIX A. FUEL MATRIX DEGRADATION MODELING IN FORTRAN

INTRODUCTION

The Fuel Matrix Degradation (FMD) model is a reactive/transport model that implements electrochemical reactions to predict the dissolution of spent uranium oxide nuclear fuel. It is designed to provide the radionuclide source term for performance assessment codes (e.g., PFLOTRAN). Surrogate models for the FMD process model provide significant improvements in speed for high throughput modeling of various reference cases. These models are trained on data produced by the FMD process model; bootstrapping the surrogate models to the data on which they were trained. If the underlying process model is updated, due to new experimental data, the surrogate models necessarily have to be retrained. The current implementation of the FMD process model is slow, has convergence issues, and produces discontinuities in outputted fluxes. Therefore, improving the efficiency and reliability of the FMD process model is crucial. Efforts this year focused on understanding the control parameters within the models, implementation of an adaptive timestep routine, and studying the interplay between the various reactions included within the model. Herein we will describe current and on-going efforts to identify computational bottlenecks in the FMD process model and the refactoring process. A brief overview of the model will be provided including the list of the tracked species, the reactions included at the fuel and steel interfaces, and the bulk aqueous reactions. The main bottleneck in the code will be identified and a new algorithm will be proposed. Lastly, the in-development implementation of the model in Fortran and current results will be discussed.

FMD PROCESS MODEL

We will provide a brief summary of the key features of the FMD model, however the model has been described in much greater detail in several references (Jerden, et al., 2012; Jerden, et al., 2018) The FMD model seeks to predict the degradation rate of UO_2 using a Mixed Potential Model approach, originally based on the Canadian Mixed Potential Model of Shoesmith, King, and Kolar (Shoesmith, et al., 2003). A conceptual image of the overall reaction scheme is shown in Fig. A-1.

The Mixed Potential Approach works on two assumptions: (1) electrochemical reactions can be broken into partial oxidation/reduction reactions, and (2) no net accumulation of charge is observed (Wagner and Traud, 1938). The 1-dimensional physical model is broken into 3 regions: (1) fuel surface, (2) solution or aqueous chemistry, and (3) steel surface. The main oxidizing agent (H_2O_2), which promotes fuel dissolution, is generated via an analytical expression for a PNNL radiolysis model (Buck, et al., 2013). All diffusion and reaction rate constant incorporate an Arrhenius temperature dependence.

The fuel and steel interfacial regions are critically important and are where the Fig. electrochemical reactions are observed. Hydrogen generation and oxidation reactions occur at these interfaces and work to slow the fuel dissolution rate. The so-called “hydrogen effect” has been identified in numerous experimental and modeling efforts (Broczkowski, et al., 2005; Trummer and Jonsson, 2010; Jerden et al., 2015). Because hydrogen has been shown to reduce the rate of spent fuel dissolution, a simple approach would be to ignore the effect and provide a conservative estimate for the radionuclide source term in performance assessment. However, this may not be realistic for certain repository environments.

Reactions at the fuel surface can be catalyzed by noble metal particles typically referred to as the ϵ phase. This phenomenon has been observed in electrochemical experiments which showed that dissolved H_2 lowered the spent fuel corrosion potential (i.e., E_{corr} , the voltage at which the anodic and cathodic reaction rates are identical) more in samples that contained the ϵ phase (Broczkowski et al., 2005; Shoesmith, 2008).

In addition to the electrochemical reactions, solution chemistry is included which accounts for typical reactions observed involving aqueous ions (e.g., Fe^{2+} , UO_2^{2+}). A list of the solution chemistry reactions and their rate constants included in the model are shown in Tab. A-1. UO_2^{2+} can also precipitate at the fuel surface in the form of schoepite and metaschoepite. This solid phase protects the fuel surface from oxidants and reductants by slowing the diffusion rate according to a user-inputted porosity parameter.

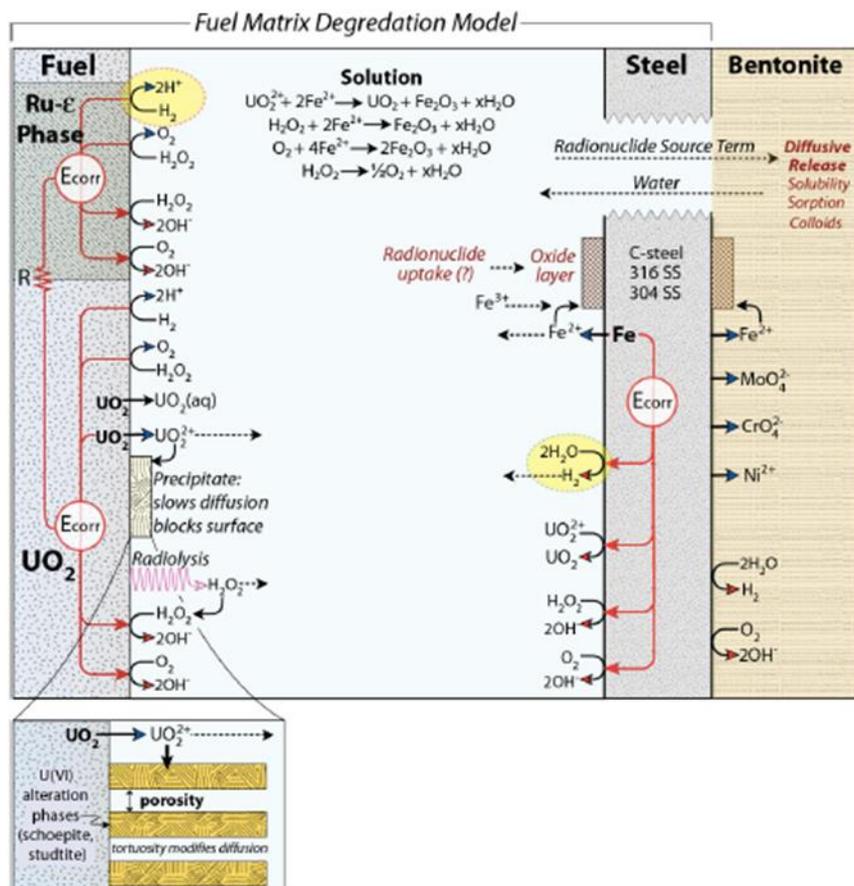


Fig. A-1 Schematic depicting the physical model described by the FMD model. Three regions are included: fuel surface, bulk solution, and steel surface. Electrochemical reactions are observed in the fuel and steel regions, where noble metal particles (Ru-ε) can catalyze these reactions at the fuel surface. Fuel (UO_2) dissolves, via chemical dissolution or oxidation, to form UO_2^{2+} which diffuses to the breached steel surface and becomes the radionuclide source term. UO_2^{2+} can also reprecipitate in the form of schoepite, forming a porous barrier and slowing diffusion to the fuel surface.

The bulk solution reactions depicted in Fig. A-1 are implemented by discretizing the system in space and time and solving the partial differential equations using a nonlinear solver. In the Mixed Potential Model approach, the electrochemical reactions at the fuel and steel surface are implemented by calculating the mass loss according to the current density as described by Faradays Law.

$$\frac{dFuel}{dt} = \frac{i_{corr}^{fuel} MW_{fuel}}{nF} \quad (A-1)$$

where i_{corr}^{fuel} is the corrosion current density, MW_{fuel} is the molecular weight of the fuel, n is the number of transferred electrons, and F is Faraday's constant. A current density for each reaction can be defined and is shown in Tab. A-2. A particular challenge exists where the current densities depend on both the concentrations and the corrosion potential for a particular surface ($E_{corr}^{fuel} / E_{corr}^{steel}$). In the current implementation these unknowns are solved sequentially; first solving for the corrosion potential with a given concentration, and then solving for the new concentration given the calculated corrosion potential.

As mentioned above, H_2O_2 is the primary oxidant included in the FMD model. This species is produced via alpha radiolysis. Radiolysis is incorporated into the FMD model via a straightforward analytical function which describes the production of H_2O_2 as a function of space and time (Buck, et al., 2013). It is given by

$$\text{Molar Yield of } H_2O_2(x, t) = [G_{H_2O_2}](H_2, O_2) * [\text{Dose Rate}](x, t) * g(x) \quad (\text{A-2})$$

where $g(x)$ is a geometrical factor that alters the diffusion of aqueous species through the tortuosity of precipitated UO_2 .

Tab. A-1 Bulk solution/aqueous chemistry reactions included in the FMD model. For each reaction k is the reaction rate constant and dT is a temperature factor given by $dT = \left(\frac{1}{298} - \frac{1}{T}\right) / R$ where T is the current temperature and R is the gas constant.

| Reaction | Rate Law | Rate Constant |
|--|--|---|
| $UO_2^{2+} + 2OH^- + H_2O \rightarrow UO_3 \cdot 2H_2O$ | $\frac{\partial UO_2^{2+}}{\partial t} = -k_1[UO_2^{2+}]^5$ $\frac{\partial UO_3 \cdot 2H_2O}{\partial t} = k_1[UO_2^{2+}]^5$ | k_1 $= 1.0 \times 10^{-3} e^{6.0 \times 10^4 \cdot dT / (2.4 \cdot c_{sat}^{UO_2^{2+}})^4}$ |
| $UO_2^{2+} + H_2O_2 + 4H_2O \rightarrow UO_4 \cdot 4H_2O + 2H^+$ | $\frac{\partial UO_2^{2+}}{\partial t} = -k_2[H_2O_2][UO_2^{2+}]^5$ $\frac{\partial H_2O_2}{\partial t} = -k_2[H_2O_2][UO_2^{2+}]^5$ $\frac{\partial UO_4 \cdot 4H_2O}{\partial t} = k_2[H_2O_2][UO_2^{2+}]^5$ | k_2 $= 1.0 \times 10^{-6} e^{6.0 \times 10^4 \cdot dT / (2.4 \cdot c_{sat}^{UO_2^{2+}})^4}$ |
| $UO_2(CO_3)_2^{2-} + 2OH^- + H_2O \rightarrow UO_3 \cdot 2H_2O + 2CO_3^{2-}$ | $\frac{\partial UO_2(CO_3)_2^{2-}}{\partial t} = -k_3[UO_2(CO_3)_2^{2-}]^5$ $\frac{\partial UO_3 \cdot 2H_2O}{\partial t} = k_3[UO_2(CO_3)_2^{2-}]^5$ $\frac{\partial CO_3^{2-}}{\partial t} = 2k_3[UO_2(CO_3)_2^{2-}]^5$ | k_3 $= 1.0$ $\times 10^{-4} e^{6.0 \times 10^4 \cdot dT / (2.4 \cdot c_{sat}^{UO_2(CO_3)_2^{2-}})^4}$ |
| $UO_3 \cdot 2H_2O + 2CO_3^{2-} \rightarrow UO_2(CO_3)_2^{2-} + 2OH^- + H_2O$ | $\frac{\partial UO_2(CO_3)_2^{2-}}{\partial t} = k_4[CO_3^{2-}][UO_3 \cdot 2H_2O]$ $\frac{\partial UO_3 \cdot 2H_2O}{\partial t} = -k_4[CO_3^{2-}][UO_3 \cdot 2H_2O]$ $\frac{\partial CO_3^{2-}}{\partial t} = -2k_4[CO_3^{2-}][UO_3 \cdot 2H_2O]$ | $k_4 = 8.6 \times 10^{-6} e^{6.0 \times 10^4 \cdot dT}$ |
| $H_2O_2 + 2Fe^{2+} + 4OH^- \rightarrow 3H_2O + Fe_2O_3$ | $\frac{\partial H_2O_2}{\partial t} = -k_5[H_2O_2][Fe^{2+}]$ $\frac{\partial Fe^{2+}}{\partial t} = -2k_5[H_2O_2][Fe^{2+}]$ | $k_5 = 1.0 \times 10^{-3} e^{4.2 \times 10^4 \cdot dT}$ |
| $O_2 + 4Fe^{2+} + 8OH^- \rightarrow 4H_2O + 2Fe_2O_3$ | $\frac{\partial Fe^{2+}}{\partial t} = -4k_6[Fe^{2+}][O_2]$ $\frac{\partial O_2}{\partial t} = -k_6[Fe^{2+}][O_2]$ | $k_6 = 1.0 \times 10^{-3} e^{6.0 \times 10^4 \cdot dT}$ |

| Reaction | Rate Law | Rate Constant |
|---|---|--|
| $UO_2^{2+} + 2Fe^{2+} + 6OH^- \rightarrow UO_{2,(aq)} + 3H_2O + Fe_2O_3$ | $\frac{\partial UO_2^{2+}}{\partial t} = -k_7[UO_2^{2+}][Fe^{2+}]$ $\frac{\partial Fe^{2+}}{\partial t} = -2k_7[UO_2^{2+}][Fe^{2+}]$ $\frac{\partial UO_{2,(aq)}}{\partial t} = k_7[UO_2^{2+}][Fe^{2+}]$ | $k_7 = 1.0 \times 10^{-2} e^{6.0 \times 10^4 \cdot dT}$ |
| $UO_2(CO_3)_2^{2-} + 2Fe^{2+} + 6OH^- \rightarrow UO_{2,(aq)} + 2CO_3^{2-} + 3H_2O + Fe_2O_3$ | $\frac{\partial UO_2(CO_3)_2^{2-}}{\partial t} = -k_8[UO_2(CO_3)_2^{2-}][Fe^{2+}]$ $\frac{\partial CO_3^{2-}}{\partial t} = 2k_8[UO_2(CO_3)_2^{2-}][Fe^{2+}]$ $\frac{\partial Fe^{2+}}{\partial t} = -2k_8[UO_2(CO_3)_2^{2-}][Fe^{2+}]$ $\frac{\partial UO_{2,(aq)}}{\partial t} = k_8[UO_2(CO_3)_2^{2-}][Fe^{2+}]$ | $k_8 = 1.0 \times 10^{-3} e^{6.0 \times 10^4 \cdot dT}$ |
| $H_2O_2 \rightarrow H_2O + \frac{1}{2}O_2$ | $\frac{\partial H_2O_2}{\partial t} = -k_9[H_2O_2]$ $\frac{\partial O_2}{\partial t} = \frac{1}{2}k_9[H_2O_2]$ | $k_9 = 4.5 \times 10^{-7} e^{6.0 \times 10^4 \cdot dT}$ |
| $UO_{2,(s)} \rightarrow UO_{2,(aq)}$ | $\frac{\partial UO_{2,(aq)}}{\partial t} = k_{10}[UO_{2,(s)}]$ $\frac{\partial UO_{2,(s)}}{\partial t} = -k_{10}[UO_{2,(s)}]$ | $k_{10} = 1.9 \times 10^{-12} e^{6.0 \times 10^4 \cdot dT}$ |
| $UO_{2,(aq)} \rightarrow UO_{2,(s)}$ | $\frac{\partial UO_{2,(aq)}}{\partial t} = -k_{11}[UO_{2,(aq)}]^5$ $\frac{\partial UO_{2,(s)}}{\partial t} = k_{11}[UO_{2,(aq)}]^5$ | $k_{11} = 1.0 \times 10^{-4} e^{6.0 \times 10^4 \cdot dT / (2.4 \cdot c_{sat}^{UO_{2,(aq)}})^4}$ |

Tab. A-2 Electrochemical reactions and parameters included in the FMD model where *i* is the current density for a given reaction, *k* is the half reaction rate constant, α is the charge transfer coefficient, E^0 is the standard potential, and dT is a temperature factor given by $dT = \left(\frac{1}{298} - \frac{1}{T}\right) / R$ where T is the current temperature and R is the gas constant. There are 3 surfaces (fuel, fuel/NMP, and steel) and each surface has a corrosion potential, E_{corr} , which must be determined at each timestep.

| Surface | Reactions | Equation | k | α | E^0 (V) |
|---------|--|--|---|----------|-----------|
| Fuel | $UO_2 \rightarrow UO_2^{2+} + 2e^-$ | $i^{UO_2,1} = nF\epsilon k_{UO_2,1} \exp \left[\frac{\alpha_{UO_2,1} F}{RT} (E_{corr}^{UO_2} - E_{UO_2,1}^0) \right]$ | $5.0 \times 10^{-8} \exp[6.0 \times 10^4 \cdot dT]$ | 0.96 | 0.169 |
| | $UO_2 + 2CO_3^{2-} \rightarrow UO_2CO_3^{2-} + 2e^-$ | $i^{UO_2,2} = nF\epsilon k_{UO_2,2} [CO_3^{2-}]^2 \exp \left[\frac{\alpha_{UO_2,2} F}{RT} (E_{corr}^{UO_2} - E_{UO_2,2}^0) \right]$ | $1.3 \times 10^{-8} \exp[6.0 \times 10^4 \cdot dT]$ | 0.82 | .173 |

| Surface | Reactions | Equation | k | α | E° (V) |
|-----------|---|---|--|----------|---------------|
| | $UO_2 \rightarrow UO_{2,(aq)}$ | $i^{UO_2,3} = nF\epsilon k_{UO_2,3} \exp \left[\frac{\alpha_{UO_2,3} F}{RT} (E_{corr}^{UO_2} - E_{UO_2,3}^0) \right]$ | $2.0 \times 10^{-14} \exp[6.0 \times 10^4 \cdot dT]$ | 0.00 | 0.000 |
| | $H_2 + 2OH^- \rightarrow 2H_2O + 2e^-$ | $i^{H_2} = nF\epsilon k_{H_2} [H_2] \exp \left[\frac{\alpha_{H_2} F}{RT} (E_{corr}^{UO_2} - E_{H_2}^0) \right]$ | $3.6 \times 10^{-12} \exp[6.0 \times 10^4 \cdot dT]$ | 0.50 | 0.584 |
| | $H_2O_2 + 2OH^- \rightarrow O_2 + 2H_2O + 2e^-$ | $i^{H_2O_2,1} = nF\epsilon k_{H_2O_2,1} [H_2O_2] \exp \left[\frac{\alpha_{H_2O_2,1} F}{RT} (E_{corr}^{UO_2} - E_{H_2O_2,1}^0) \right]$ | $7.4 \times 10^{-8} \exp[6.0 \times 10^4 \cdot dT]$ | 0.41 | -0.098 |
| | $H_2O_2 + 2e^- \rightarrow 2OH^-$ | $i^{H_2O_2,2} = nF\epsilon k_{H_2O_2,2} [H_2O_2] \exp \left[\frac{\alpha_{H_2O_2,2} F}{RT} (E_{corr}^{UO_2} - E_{H_2O_2,2}^0) \right]$ | $1.2 \times 10^{-12} \exp[6.0 \times 10^4 \cdot dT]$ | -0.41 | 0.973 |
| | $O_2 + 2H_2O + 4e^- \rightarrow 4OH^-$ | $i^{O_2} = nF\epsilon k_{O_2} [O_2] \exp \left[\frac{\alpha_{O_2} F}{RT} (E_{corr}^{UO_2} - E_{O_2}^0) \right]$ | $1.4 \times 10^{-12} \exp[6.0 \times 10^4 \cdot dT]$ | -0.50 | 0.157 |
| Fuel, NMP | $H_2 + 2OH^- \rightarrow 2H_2O + 2e^-$ | $i^{H_2} = nF\epsilon k_{H_2} [H_2] \exp \left[\frac{\alpha_{H_2} F}{RT} (E_{corr}^{NMP} - E_{H_2}^0) \right]$ | $5.0 \exp[6.0 \times 10^4 \cdot dT]$ | 0.50 | 0.584 |
| | $H_2O_2 + 2e^- \rightarrow 2OH^-$ | $i^{H_2O_2,1} = nF\epsilon k_{H_2O_2,1} [H_2O_2] \exp \left[\frac{\alpha_{H_2O_2,1} F}{RT} (E_{corr}^{NMP} - E_{H_2O_2,1}^0) \right]$ | $1.2 \times 10^{-12} \exp[6.0 \times 10^4 \cdot dT]$ | -0.40 | 0.973 |
| | $H_2O_2 + 2OH^- \rightarrow O_2 + 2H_2O + 2e^-$ | $i^{H_2O_2,2} = nF\epsilon k_{H_2O_2,2} [H_2O_2] \exp \left[\frac{\alpha_{H_2O_2,2} F}{RT} (E_{corr}^{UO_2} - E_{H_2O_2,2}^0) \right]$ | $5.0 \times 10^{-12} \exp[6.0 \times 10^4 \cdot dT]$ | 0.41 | -0.098 |
| | $O_2 + 2H_2O + 4e^- \rightarrow 4OH^-$ | $i^{O_2} = nF\epsilon k_{O_2} [O_2] \exp \left[\frac{\alpha_{O_2} F}{RT} (E_{corr}^{UO_2} - E_{O_2}^0) \right]$ | $1.4 \times 10^{-12} \exp[6.0 \times 10^4 \cdot dT]$ | -0.50 | 0.157 |
| Steel | $Fe \rightarrow Fe^{2+} + 2e^-$ | $i^{steel} = nF\epsilon k_{steel} \exp \left[\frac{\alpha_{steel} F}{RT} (E_{corr}^{steel} - E_{steel}^0) \right]$ | $1.0 \times 10^{-6} \exp[1.0 \times 10^4 \cdot dT]$ | 0.20 | 0.203 |
| | $2H_2O + 2e^- \rightarrow H_2 + 2OH^-$ | $i^{H_2O} = nF\epsilon k_{H_2O} \exp \left[\frac{\alpha_{H_2O} F}{RT} (E_{corr}^{steel} - E_{H_2O}^0) \right]$ | $1.0 \times 10^{-6} \exp[1.0 \times 10^4 \cdot dT]$ | -0.48 | -1.072 |

FMD IMPLEMENTATION

Currently the most regularly updated and used implementation of the FMD model is a MATLAB version. While a stand-alone Fortran version exists (Jerden, et al., 2015), it is a translation of the MATLAB version into a new programming language. The MATLAB implementation exists as a main program and several subroutines and functions that make use of multiple built-in MATLAB functionalities. A flow chart for the MATLAB version of the FMD model is depicted in Fig. A-2. A main program (AMP_main) launches the simulation and drives the loop over time points. A function (`AMP_solve`) is called at each time point which runs the standard Newton Raphson (NR) nonlinear solver to update the concentrations. For each NR iteration the Jacobian and Residual are calculated in three stages using AMP_reactFuel, AMP_reactCani, AMP_reactBulk which calculates the components to each at the fuel surface, canister surface, and bulk reaction regions. However, the concentration of several components are dependent upon the corrosion potential, therefore for each new concentration solve the corrosion potential at both surfaces must be solved using a separate NR process.

The number of NR iterations needed to solve for the new concentrations at each time point, from a representative FMD simulation using the MATLAB implementation is shown in Fig. A-3. In general, 5-10 NR iterations are needed at each time point. Now we show that for *each* of these iterations the number of iterations needed to solve for the corrosion potentials at the fuel and steel surface in Fig. A-4. Note that the number of iterations needed solve for the corrosion potentials is often the same, we therefore color each point in Fig. A-4 to indicate the number of identical iterations was needed at each time point. It is clear from this analysis that significant computing time is spent sequentially iterating over dependent processes (new concentrations and corrosion potentials). In fact, a timing test was performed and observed that ~30% of the computational time is spent solving for the corrosion potentials. *We hypothesize that we can couple the NR solves for the new concentration and corrosion potentials.* Similar to the evolution of reactive-transport modeling from sequential reaction and transport iterations to the global implicit approach where a fully-coupled nonlinear system of equations for both reaction and transport are included in the Jacobian, we believe that a significant speed up will be observed by reducing the overall number of iterations needed.

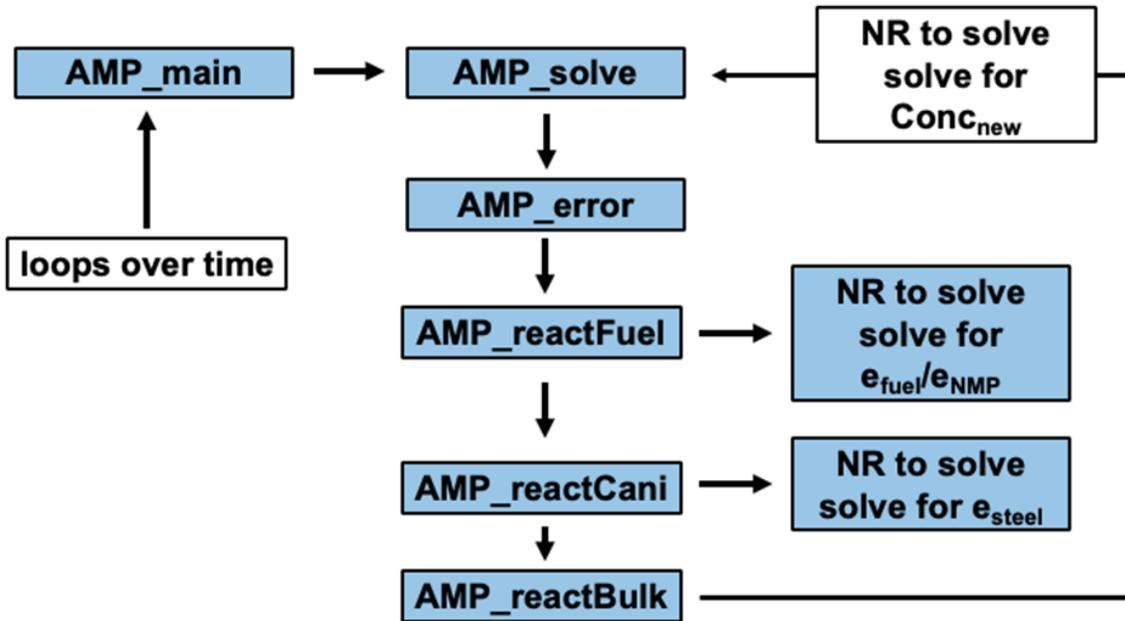


Fig. A-2 Flow chart for the FMD model implemented in MATLAB. AMP_main starts the simulation and iterates over time. AMP_solve runs the standard Newton Raphson (NR) iteration to solve for the new concentrations. Within each NR iteration, the corrosion potential at the fuel surface (AMP_reactFuel) and canister surface (AMP_reactCani) are solved iteratively using a separate NR process.

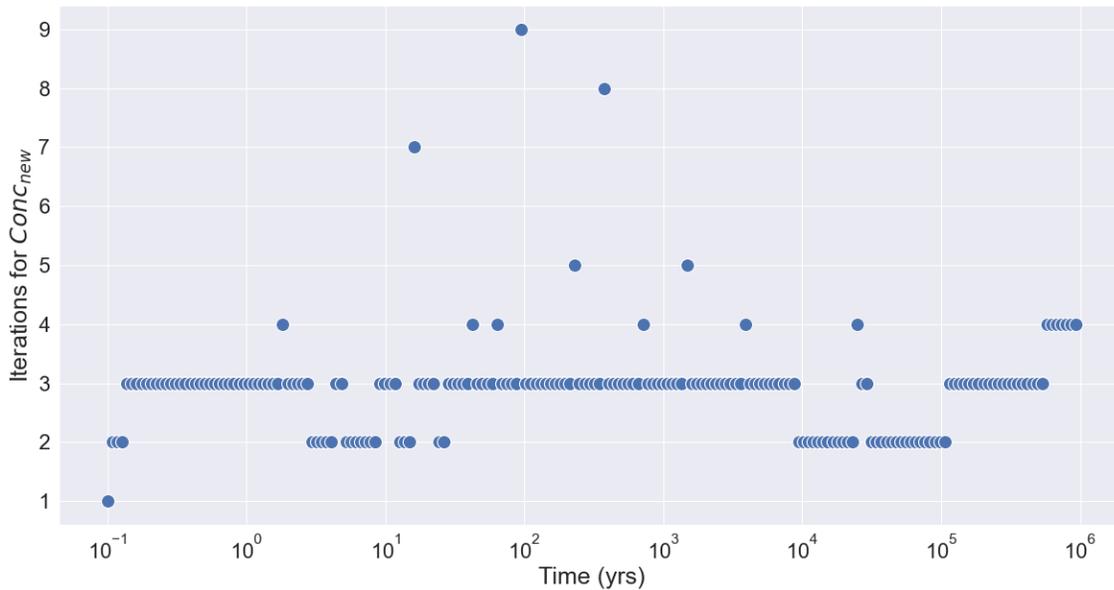


Fig. A-3 The number of Newton Raphson (NR) iterations needed at each time point throughout the simulation to solve for the new concentrations.

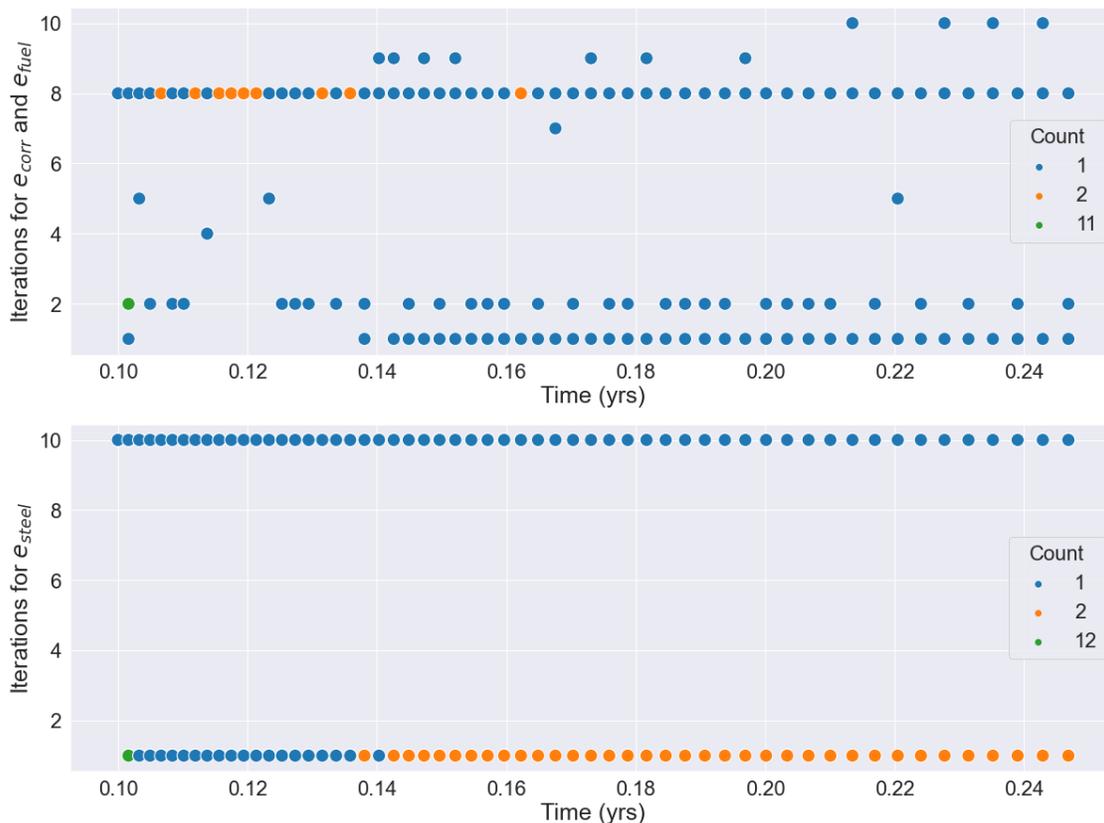


Fig. A-4 Number of Newton Raphson (NR) iterations needed to solve for the corrosion potential at the fuel surface (top) and canister surface (bottom). Note that for each time point a NR solve is performed to update the concentrations, and for each of those NR iterations, the number of iterations shown are performed. The identical number of iterations is often needed and indicated by the coloring of each point.

We will briefly describe the hypothetical approach that will be pursued in future efforts. Imagine a 1D reactive-transport system that incorporates 3 species (A, B, and C). The standard residual is given by

$$f(c^{k+1,p}) = \frac{c^{k+1,p} - c^k}{\Delta t} + L(c^{k+1,p}) - R(c^{k+1,p}) \quad (\text{A-3})$$

where c is the concentration of a given species, L is the diffusion operator, and R is the kinetic rate expression. We will now introduce a fictitious function for the fuel corrosion potential ($g(E_{corr}^{fuel})$). In the interest of simplicity, we introduce a single function, however the FMD model currently has two corrosion potential functions. Assuming the corrosion potential function only depends on the corrosion potential and the concentration of A, (E_{corr}^{fuel}, c^A), the residual is transformed to

$$f(c^{k+1,p}) = \frac{c^{k+1,p} - c^k}{\Delta t} + L(c^{k+1,p}) - R(c^{k+1,p}) + g(E_{corr}^{fuel}, c^A). \quad (\text{A-4})$$

The derivative can then be computed to determine the Jacobian. New concentrations and corrosion potentials are updated using the standard NR approach. In matrix form this has the effect of adding an additional row (new equation) and additional column (new unknown) to the usual submatrices. However,

we note that this *only* applies to the grid cells near the fuel surface. Moreover, the corrosion potential variable (E_{corr}^{fuel}) appears in no other functions, therefore most of the new entries in the Jacobian will be 0. And lastly, in this example, we have assumed the corrosion potential function is solely a function of the corrosion potential (E_{corr}^{fuel}) and the concentration of species A (c^A), therefore there will only be nonzero entries in the columns in which the derivatives with respect to these variables are performed. Note, in practice, there could be multiple corrosion potential functions which could depend on the concentrations of several species.

CURRENT EFFORTS

Given the identification of the computational bottleneck, our approach is to refactor the FMD algorithm, with the philosophy of creating a stand-alone Fortran version that could be (1) more easily used and updated by external collaborators, and (2) be modified for other potential corrosion applications. In this vein we seek to develop, from scratch, a 1D reactive-transport code that implements the FMD model described above. What follows describes that current effort.

Diffusion is implemented by discretizing the system in space and time using the finite volume approach. Imagine a 1D column with 3 grid cells and 1 diffusing species. The resulting discretized governing equation is given by (ignoring advection, kinetic reaction, and the source term)

$$V \frac{(\phi C)_i^{k+1} - (\phi C)_i^k}{\Delta t} - \left(\phi AD \frac{C_i - C_{i-1}}{\Delta x} \right)_{i-\frac{1}{2}}^{k+1} + \left(\phi AD \frac{C_{i+1} - C_i}{\Delta x} \right)_{i+\frac{1}{2}}^{k+1} = 0 = f(C^{k+1,i}) \quad (\text{A-5})$$

where k and k+1 are the current and new time step, V is the volume, ϕ is the porosity, C_i is the concentration in the ith grid cell, Δt is the time step, Δx is the distance between grid point centers, A is the area, and D is the diffusion coefficient. In this 1D case we will assume Δy and Δz are 1 and therefore $V = \Delta x$ and $A = 1$. We will also assume a *full* grid cell exists at the boundaries. Concentrations are updated using the NR approach while a linear solver is used to solve the equation

$$J\delta C = -f(C^{k+1,i}) \quad (\text{A-6})$$

In matrix form this looks like

$$\begin{bmatrix} \frac{V\phi}{\Delta t} - \frac{\phi 2AD}{\Delta x} & \frac{\phi AD}{\Delta x} & 0 \\ \frac{\phi AD}{\Delta x} & \frac{V\phi}{\Delta t} - \frac{\phi 2AD}{\Delta x} & \frac{\phi AD}{\Delta x} \\ 0 & \frac{\phi AD}{\Delta x} & \frac{V\phi}{\Delta t} - \frac{\phi 2AD}{\Delta x} \end{bmatrix} \begin{bmatrix} \delta C_1 \\ \delta C_2 \\ \delta C_3 \end{bmatrix} = \begin{bmatrix} -f(C_1^{k+1}) \\ -f(C_2^{k+1}) \\ -f(C_3^{k+1}) \end{bmatrix} \quad (\text{A-7})$$

Adding reactivity to this matrix transforms this matrix into a large $(N_c * N_{\text{grid}}) \times (N_c * N_{\text{grid}})$ where N_c is the number of components and N_{grid} is the number of grid cells. The current version of the Fortran code is version controlled and maintained in a gitlab repository.

EXPLORATION OF THE ROLE OF THE c_{sat} PARAMETERS

Several of the reactions in Tab. A-1 include a c_{sat} parameter. As we will lay out below, this parameter works to turn reactions off as the temperature decreases. For this case study, we will consider reaction 2 from Tab. A-1 (i.e., $UO_2^{2+} + H_2O_2 + 4H_2O \rightarrow UO_4 \cdot 4H_2O + 2H^+$) meaning this, and only this reaction, will be turned on. A 5-grid cell system was considered with the initial concentrations of each species shown in Fig. A-5. In this system, the size of each grid cell is the same. The boundary condition at both the left and right boundaries was 0.01 M for all species. Diffusion was turned on, however $UO_4 \cdot 4H_2O$ is a solid and does not diffuse in the model.

| | | | | |
|--|-----|-----|-----|-----|
| $[UO_2^{2+}] = [H_2O_2] =$ $[UO_4 \cdot 4H_2O] = .01 \text{ M}$ | 0 M | 0 M | 0 M | 0 M |
|--|-----|-----|-----|-----|

Fig. A-5 5 grid cell considered for the reaction $UO_2^{2+} + H_2O_2 + 4H_2O \rightarrow UO_4 \cdot 4H_2O + 2H^+$ in which UO_2^{2+} , H_2O_2 , and $UO_4 \cdot 4H_2O$ are the species that are tracked. The initial concentration for each species was 0 M for all grid cells except for the left most grid cell where the initial concentration was 0.01 M. The size of each cell is identical.

The concentration of UO_2^{2+} as a function of time is shown as a log-log plot in Fig. A-6. Note that diffusion occurs at the left and right boundary and therefore, within a short period of time, the system is symmetric (i.e., grid cell 1 = grid cell 5). We note that the concentration of UO_2^{2+} reaches an initial equilibrium and the concentration remains relatively unchanged for ~100 years. At this point the concentration begins to increase towards that boundary condition concentration (i.e., 0.01 M) at ~1000 years.

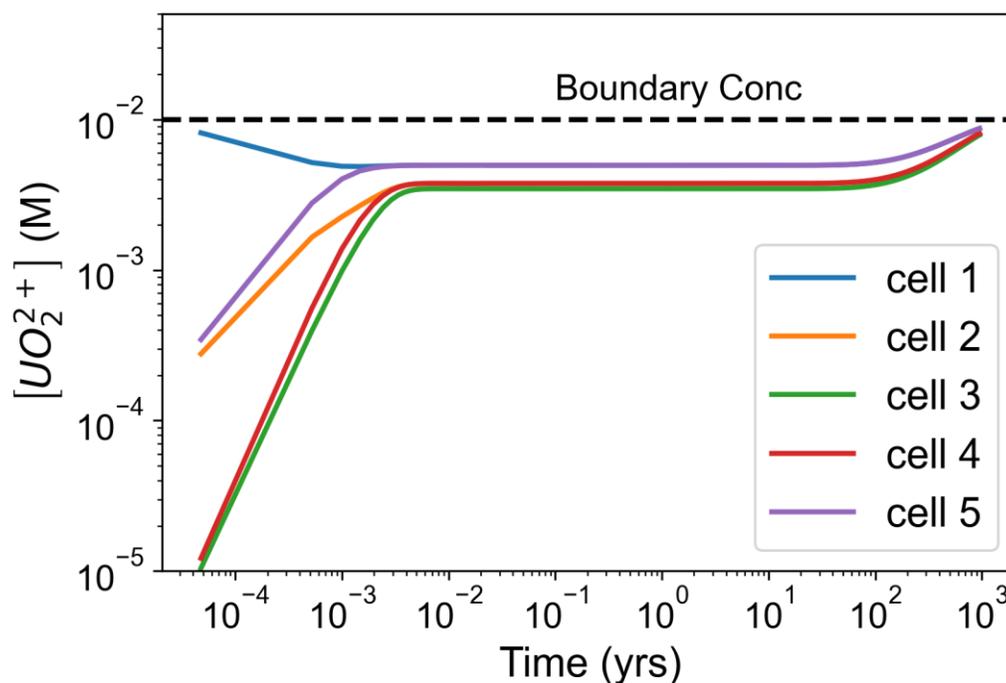


Fig. A-6 Log-log plot of the concentration (M) of UO_2^{2+} as a function of time for the 5 grid cells considered. The boundary condition (0.01 M) is indicated.

In an effort to understand this nearly 100 year equilibrium we looked at the reaction rate constant (i.e., $k_2 = 1.0 \times 10^{-6} e^{6.0 \times 10^4 \cdot dT / (2.4 \cdot c_{sat}^{UO_2^{2+}})}$) and the $c_{sat}^{UO_2^{2+}}$ parameter ($c_{sat}^{UO_2^{2+}} = 3.2 \times 10^{-2} \cdot e^{-3.0 \times 10^4 \cdot dT}$). Note that $dT = (\frac{1}{298} - \frac{1}{T}) / R$ is an Arrhenius temperature factor. The reaction rate constant and $c_{sat}^{UO_2^{2+}}$ as a function of time are shown in Fig. A-7. The temperature as a function of time is shown as an inset. Here we observe that the reaction rate constant is constant up until 100 years. At this point the temperature begins to decrease which leads to an increase in $c_{sat}^{UO_2^{2+}}$. Ultimately $c_{sat}^{UO_2^{2+}}$ increases by a factor of 2.3 from its initial value to its final value at the minimum temperature observed within this time period. However, the reaction rate constant has a $(c_{sat}^{UO_2^{2+}})^4$ dependence meaning the rate constant decreases by a factor of 31. Therefore, the observed result that the concentration of UO_2^{2+} can go unchanged for 100 years can be explained by all the concentration that diffuses in from the boundary is immediately converted to $UO_4 \cdot 4H_2O$, however as the temperature decreases the reaction rate constant decreases rapidly. At this point, the rate of diffusion is greater than the rate of the reaction and the concentration of UO_2^{2+} begins to head toward the boundary condition. In this regard $c_{sat}^{UO_2^{2+}}$ works to, effectively, turn reactions off as the temperature of the system decreases at long times. Note that the temperature of the system is a simple analytical function however in the full FMD model UO_2^{2+} is produced due to corrosion at the fuel surface. Therefore, it is theoretical possible for the reaction rate to remain high despite a decreasing temperature if the concentration of UO_2^{2+} is large.

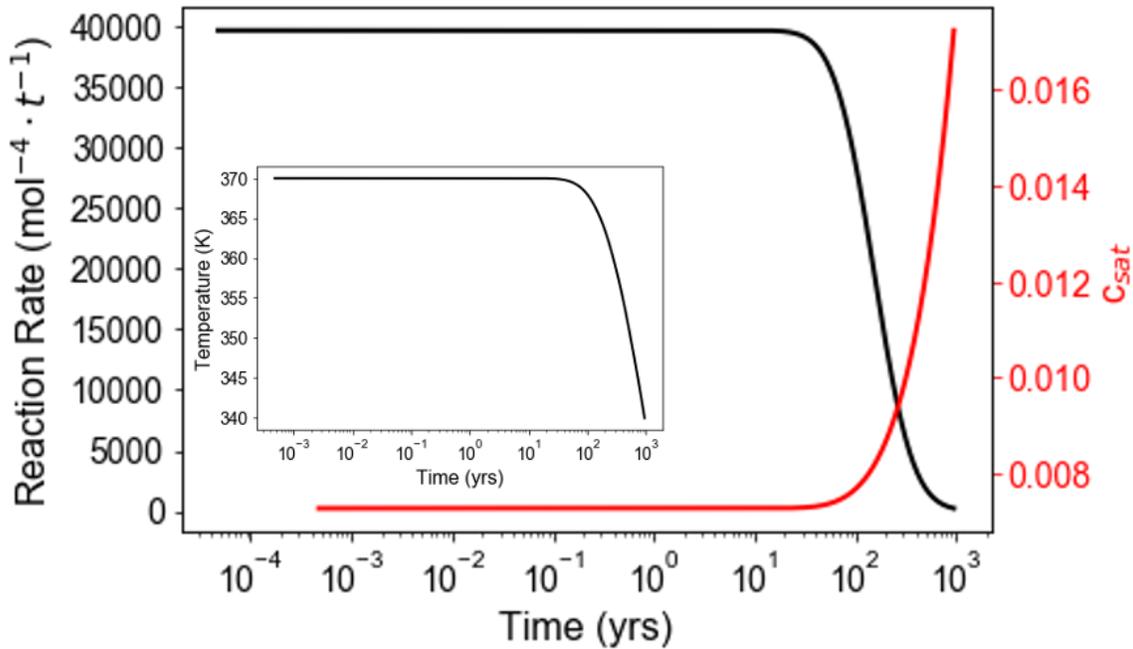


Fig. A-7 The reaction rate constant (black, left axis) of reaction 2 in Tab A-1 and the $c_{sat}^{UO_2^{2+}}$ parameter (red, right axis) as a function of time shown as a semi-log plot. The temperature as a function of time is shown in the inset.

ADAPTIVE TIMESTEP ROUTINE

The results in the above analysis on the $c_{sat}^{UO_2^{2+}}$ parameter indicate that long simulation times would be needed for reactions to reach an equilibrium. However, initial conditions are typically far from equilibrium and require short timestep periods ($< 1s$). The number of simulation steps required to reach > 1000 years would be computationally intractable using the short initial timestep throughout the entire simulation. In order to combat this problem, we implemented an adaptive time stepping routine. We took a relatively conservative approach and implemented a sliding scale for time step increases which is controlled by the number of NR iterations needed to update the concentrations. In our algorithm the timestep is increased by 5, 10, 15, and 20% if 4, 3, 2, and 1 NR iterations are required to update the concentrations. Similarly, for 8 and 9 NR iterations the timestep is decreased by 25, and 50%. If the number of NR iterations hits a user inputted max value, the timestep is cut in half, and the simulation step is restarted (i.e., the Jacobian and residual are re-calculated). Additional parameters that can be altered include the first simulation step at which the timestep is altered and the frequency in which the timestep is altered. In general, we have found that starting the timestep adjustments after the 1000th simulation step allows the system to equilibrate from the initial conditions. Similarly, altering the timestep every 1000 simulation steps has been found to be a stable setting.

REACTION RATE EXPRESSIONS

It is interesting to consider the reaction rate expressions and how they can affect the shape of the time evolution curves for various concentrations. Many of the reactions are 5th order with respect to a certain concentration. For example, reaction 3 and 4 from Tab. A-1 are the reverse of each other, however reaction 3 has a rate constant that is, initially, 100X that of reaction 4. This is due, in part, to the constant in front of the exponential but also due to the $c_{sat}^{UO_2(CO_3)_2^{2-}}$ parameter that appears in the rate constant for reaction 3 but not reaction 4. However, the rate expression has a 5th order dependence on $[UO_2(CO_3)_2^{2-}]$ (i.e., $\frac{\partial UO_2(CO_3)_2^{2-}}{\partial t} \propto [UO_2(CO_3)_2^{2-}]^5$). We investigated this effect by considering a system containing 3 grid cells of equal length. The first 4 reactions from Tab. A-1 were included. We ran two test cases, one in which the initial concentration of all species was 0.1 M and another where the initial concentration of all species was 0.01 M. The $[UO_2(CO_3)_2^{2-}]$ and $[CO_3^{2-}]$ as a function of time for both cases (lines = 0.1 M initial, points = 0.01 M initial) is shown in Fig. A-8. Note that the system is symmetric and therefore grid cells 1 and 3 are equal and therefore their curves are on top of each other.

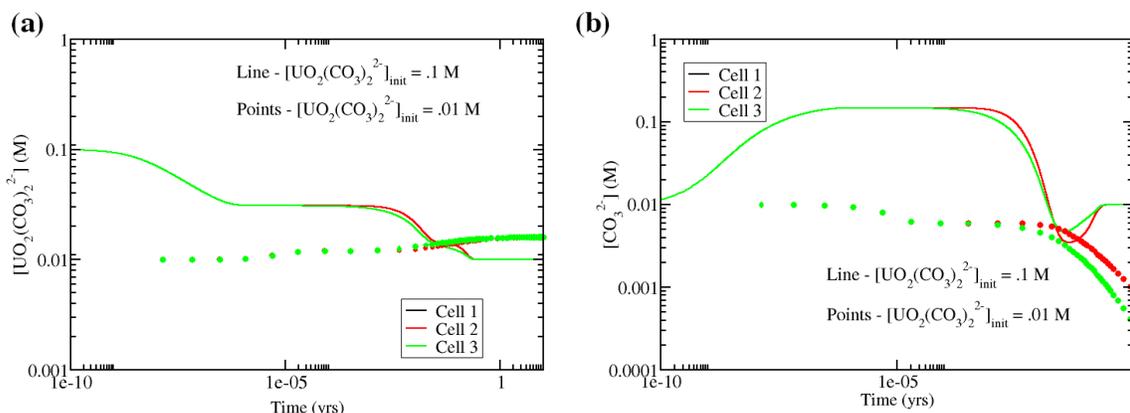


Fig. A-8 (a) Concentrations (M) of $[UO_2(CO_3)_2]^{2-}$ and (b) $[CO_3]^{2-}$ as a function of time for a 3 grid cell simulation that tracks the first 4 reactions shown in Tab. A-1. Two simulations are shown in which the initial $[UO_2(CO_3)_2]^{2-}$ concentration was 0.1 M (lines) and 0.01 M (points).

Here we observe two competing trends. When the initial concentration of $UO_2(CO_3)_2^{2-}$ is .1 M, reaction 3 is faster as CO_3^{2-} is initially created and $UO_2(CO_3)_2^{2-}$ is initially consumed. Eventually this reverses, reaction 4 becomes faster and CO_3^{2-} is over-depleted and eventually plateaus at the boundary condition (.01 M). $UO_2(CO_3)_2^{2-}$ is always consumed due to both reaction 3 and the boundary condition (i.e., concentration flows out of the cell). However, when the initial concentration of $UO_2(CO_3)_2^{2-}$ is .01 M the opposite is observed. Reaction 4 is clearly faster initially and CO_3^{2-} is destroyed. Note that if the simulation is run long enough CO_3^{2-} would revert to the boundary condition here as well. Here it is clear that the 5th order reaction dependence can have dramatic effects on the evolution of the system. Despite the reaction rate constant for reaction 3 being orders of magnitude larger than reaction 4, the reaction can proceed slower if the concentration of $UO_2(CO_3)_2^{2-}$ is low. We note that significant changes in the results can be observed based on varying initial conditions and boundary conditions and how that affects the overall evolution of the system can be non-intuitive and difficult to predict. This is a potential area of improvement in the FMD model where a less harsh reaction rate expression could be used to turn reactions on and off. We emphasize that the FMD model has compared well with experimental measures previously and therefore this is being suggested not to improve the performance of the model but to make the model more intuitive.

CONCLUSION

The computational bottleneck in the MATLAB-implemented FMD model has been identified using extensive timing studies. Significantly redundant NR iterations are being performed that can likely be reduced if the concentrations and corrosion potentials are solved for simultaneously. We have proposed a potential algorithm that could be used to achieve this and are currently working on implementation. We have shown in great detail how various reactions are controlled by certain parameters within the model and how initial and boundary conditions can significantly change the evolution of the system. Our current stand-alone 1D reactive-transport code includes all bulk solution reactions, implements an adaptive time stepping routine, and is maintained via a gitlab repository with full version control. This effort is crucial for increasing the efficiency of surrogate model training on the FMD model.

REFERENCES

- Broczkowski, M.E., Noel, J.J., Shoesmith, D.W. (2005), The Inhibiting Effects of Hydrogen on the Corrosion of Uranium Dioxide Under Nuclear Waste Disposal Conditions, *J. Nucl. Mater.*, 346:16-23.
- Buck, E., Jerden, J., Ebert, W., Wittman, R. (2013), Couple the Mixed Potential and Radiolysis Models for Used Fuel Degradation, FCRD-UFD-2013-000290, Argonne National Laboratory.
- Jerden, J., Frey, K., Cruse, T., Ebert, W. (2012), Waste Form Degradation Model Status Report: Electrochemical Model for Used Fuel Matrix Degradation Rate, FCRD-UFD-2012-000169, Argonne National Laboratory.
- Jerden, J., Copple, J.M., Frey, K.E., Ebert, W. (2015), Mixed Potential Model for Used Fuel Dissolution – Fortran Code, FCRD-UFD-2015-000159, Argonne National Laboratory.
- Jerden, J., Gattu, V.K., Ebert, W. (2018), Update on Validation and Incorporation of a New Steel Corrosion Module into Fuel Matrix Degradation Model, M4SF-18AN010301017, Argonne National Laboratory.
- Patankar, S.V. (1980). *Numerical Heat Transfer and Fluid Flow*, Hemisphere Publishing Corporation.
- Shoesmith, D.W., Kolar, M., King, F. (2003), A Mixed-Potential Model to Predict Fuel (Uranium Dioxide) Corrosion within a Failed Nuclear Waste Container, *Corrosion*, 59:802-816.
- Shoesmith, D., The Role of Dissolved Hydrogen on the Corrosion/Dissolution of Spent Nuclear Fuel, Nuclear Waste Management Organization, Toronto, Canada, TR-2008-19, November 2008.
- Trummer, M., Jonsson, M. (2010), Resolving the H₂ Effect on Radiation Induced Dissolution of UO₂-based Spent Nuclear Fuel, *J. Nucl. Mater.*, 396:163-169.
- Wagner, C., Traud, W. (1938), *Elektrochem.*, 44:391-454.

(This page is intentionally blank.)

Appendix B. SURROGATE MODELING OF THE FUEL MATRIX DEGRADATION (FMD) PROCESS MODEL

APPENDIX B. SURROGATE MODELING OF THE FUEL MATRIX DEGRADATION (FMD) PROCESS MODEL

INTRODUCTION

In model simulations of deep geologic repositories, UO₂ fuel matrix degradation typically begins as soon as the waste package (WP) breaches and groundwater contacts the fuel surface. The initial degradation rate depends on the timing of these events, burnup of the fuel, temperature, and concentrations of dissolved reactants.

Estimating the initial rate of degradation is fairly straightforward, but as UO₂ corrosion products precipitate on the fuel surface and the movement of dissolved species between the fuel surface and environment is impeded by the precipitated solids, the rate is more difficult to quantify. At that point, calculating the degradation rate becomes a reactive-transport problem in which a large number of equations must be solved by iteration for a large number of grid cells at each time step. The consequence is that repository simulations, which are already expensive, become much more expensive, especially when hundreds or thousands of WPs breach.

The Fuel Matrix Degradation (FMD) model is the process model of the Spent Fuel and Waste Science and Technology (SFWST) campaign of the US Department of Energy (DOE). It calculates spent fuel degradation rates as a function of radiolysis, redox reactions, electrochemical reactions, alteration layer growth, and diffusion of reactants through the alteration layer (Jerden et al. 2015). Like other similar fuel degradation process models, it is a complicated model requiring a large number of calculations and iterations at each time step.

One way to reduce the cost of repository simulations that include the FMD process model is to assume that all UO₂ in the repository (or certain sections of the repository) degrades at the same FMD process model rate [M L⁻² T⁻¹], adjusting for the time of WP breach. This way, the degradation process model only needs to be simulated once (or once for each section of the repository). A major drawback of this approach, however, is that temperature and environmental concentrations of reactants can vary widely across the repository in both space and time. Consequently, this approach will not propagate heterogeneity to fuel matrix degradation rates in the simulation. A more comprehensive discussion of the drawbacks of this homogenizing approach is provided in Mariner et al. (2018, Section 3.2.5).

To include the effects of spatial and temporal heterogeneity in variables influencing fuel matrix degradation rates in repository performance assessment (PA) simulations, a faster calculation of these rates is needed. The work in this report was initiated based on the hypothesis that surrogate models can be developed to inexpensively provide accurate degradation rates in a repository simulation for each individual breached WP in its own evolving environment at each time step.

Surrogate Modeling

A surrogate model (sometimes called meta-model, emulator, or response surface model) is an input-to-output mapping that replaces a more complicated simulation code. Once constructed, this meta-model is relatively inexpensive to evaluate so it is often used as a surrogate for the physics model in uncertainty propagation, sensitivity analysis, or optimization problems that may require thousands to millions of function evaluations (Simpson et al. 2008).

There are many different types of surrogate models, including neural networks, regression models, radial basis functions, splines, etc. One popular approach in the literature is to develop an emulator that is a stationary smooth Gaussian process (Santner et al. 2003; Rasmussen and Williams 2006). There are many good overview articles that compare various meta-model strategies. Various smoothing predictors and nonparametric regression approaches are compared elsewhere (Santner et al. 2003; Simpson et al. 2008; Storlie et al. 2009). Simpson et al. (2008) provides an excellent overview not just of various statistical meta-model methods but also approaches that use low-fidelity models as surrogates for high-fidelity models.

Three different types of surrogate models were developed for the FMD process model: a polynomial regression surrogate, an artificial neural network (ANN) surrogate, and a k-Nearest-Neighbors regression (kNNr) surrogate (Mariner et al. 2019, Appendix A). The polynomial and ANN surrogates are parametric models while the kNNr surrogate is nonparametric. In FY 2019, the ANN surrogate was found to be more accurate than the polynomial surrogate for this application; therefore, only the ANN and kNNr surrogates were pursued thereafter. The ANN surrogate utilizes a network of artificial neurons with nonlinear activation functions. The kNNr surrogate uses an advanced technique to interpolate between points in a multidimensional lookup table.

Objectives

This report is a follow-on to Mariner et al. (2020), which was a major milestone of the FMD surrogate modeling work that began in FY 2018. Two surrogate models of the FMD process model, an ANN surrogate and a k-Nearest-Neighbors regression (kNNr) surrogate, were developed and tested for speed and accuracy and were coupled with PFLOTRAN for use in repository PA simulations.

The objectives of this report are to:

- Review the motivation for developing the FMD surrogate models,
- Describe the development and use of the surrogate models,
- Quantify the accuracy and speed of the surrogate models,
- Describe the coupling of the surrogate models to PFLOTRAN,
- Demonstrate the surrogate models in a repository reference case,
- Discuss lessons learned, and
- Propose future improvements to the surrogate models.

FMD PROCESS MODEL

The FMD process model calculates spent fuel degradation rates as a function of radiolysis, alteration layer growth, and diffusion of reactants through the alteration layer. This model was developed at Argonne National Laboratory and Pacific Northwest National Laboratory.

Coded in MATLAB, the FMD process model incorporates two general models, a mixed potential model (MPM) and an analytical radiolysis model. The MPM is based on a model developed by (King and Kolar 1999; King and Kolar 2003). It simulates interfacial electrochemical reactions and reactive transport processes between the fuel surface and bulk water.

The FMD process model used in this report (and used to develop the surrogate models) is the version that aligns with version 2 of the MPM (Jerden et al. 2014). Hereafter, this version is called FMD V2. Version 2 of the MPM includes NMP catalysis of redox reactions as an additional process. FMD V2 includes radiolysis. Radiolysis generates aqueous species, e.g., H_2O_2 , that further augment fuel oxidation when the fuel surface dose rate is high.

Starting with version 3 of MPM and FMD, steel corrosion was added to provide a source of hydrogen (Jerden et al. 2018). This change is not included FMD V2. Steel corrosion is excluded from this report because the plan for GDSA Framework is to develop separate modules for WP and WF degradation and to couple them later via user options in PFLOTRAN.

The subsections below describe the features and processes of FMD V2 (Features and Processes), summarizes important assumptions and ranges of validity (Assumptions and Limitations), and identifies requirements for coupling to PFLOTRAN (Coupling Requirements).

Features and Processes

The features and processes of FMD V2 are illustrated in Fig. B-1. Features include the following:

- *UO₂ fuel surface.* The UO₂ fuel surface is mostly ($\geq 99\%$) pure UO₂ (s) and partly ($\leq 1\%$) regions of noble metal particles (NMPs) (Jerden et al. 2014).
- *Corrosion layer.* The corrosion layer is the accumulation of uranium minerals on the fuel surface resulting from chemical and electrochemical precipitation reactions. It is assumed to have 50% connected porosity and a tortuosity of 0.1 (Jerden et al. 2014).
- *Environmental water.* This nearby water introduces aqueous reactants from the environment and provides a boundary condition for the model. The environmental aqueous reactants in the model are H₂, O₂, CO₃²⁻, and Fe²⁺. Each of these species either reacts with radiolytic species or is itself radiolytically active (e.g., CO₃²⁻). In addition, O₂ is an important chemical oxidant in the model.
- *Interstitial water.* This water fills the void between the fuel surface and the environmental water. Chemical reactions, radiolysis, and diffusion in both directions occur in this water.

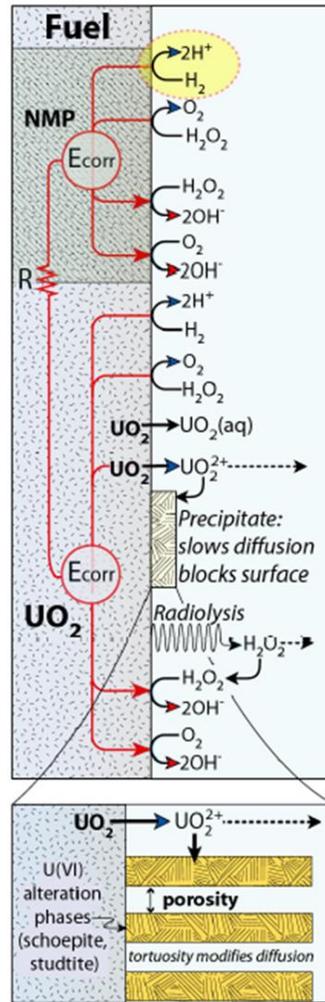


Fig. B-1 Schematic diagram of FMD features and processes implemented in PFLOTRAN and included in FMD surrogate models. Diagram adapted from Jerden et al. (2017)

The model domain depicted Fig. B-2 is one-dimensional (1D). It extends a total of 0.05 m from the fuel surface to the bulk water. This distance is divided into as many as 100 cells with increasing spatial resolution toward the cell boundaries.

The processes in FMD V2 include (Jerden et al. 2014):

- *Production of hydrogen peroxide.* Alpha radiolysis generates H₂O₂ (hydrogen peroxide) near the fuel surface. The amount of H₂O₂ generated each time step is a function of temperature, dose rate, radiolytic G-values, and the initial concentrations of aqueous species. H₂O₂ is the dominant fuel oxidant in anoxic repository environments.
- *Oxidative dissolution of the fuel matrix.* This type of dissolution oxidizes U(IV) at the surface of the UO₂(s) matrix to U(VI), releasing the U(VI) to solution. It is calculated as a function of interfacial redox reaction kinetics (based on corrosion potentials) for both pure UO₂(s) and for a fission product alloy phase referred to as the Noble Metal bearing Particles (NMP) or epsilon phase particles.

- *Chemical dissolution of the fuel matrix.* This type of dissolution releases reduced uranium as U(IV) based on solubility-driven rate calculations at the interface.
- *Other redox reactions at the fuel surface and the NMP surface.* H₂O₂, O₂, and H₂ are kinetically oxidized/reduced at the fuel surface and at NMP phase surfaces.
- *Aqueous redox reactions.* Reduction of H₂O₂, O₂, and U(VI) (aq) and oxidation of H₂ and Fe²⁺ are kinetically controlled.
- *Precipitation (and dissolution) of uranium phases.* Kinetically-driven precipitation (and dissolution) of solid U(IV) and U(VI) phases occur on the fuel surface generating a porous corrosion layer.
- *Complexation by carbonate.* Carbonate from bulk solution reacts kinetically with uranium at the fuel surface and with uranium in the corrosion layer.
- *Diffusion.* Aqueous reactants and products diffuse within the 1D domain. Diffusion is slower through the porosity of the corrosion layer.
- *Temperature dependence.* An Arrhenius temperature dependence applies to all rate constants, mineral saturation concentrations, and diffusion coefficients. A linear temperature dependence applies to standard electrochemical potentials (Jerden et al. 2012).

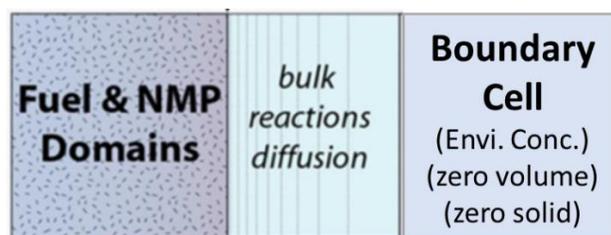


Fig. B-2 FMD process model domain

The user selected FMD V2 inputs describing fuel and environmental characteristics include:

- UO₂ fuel properties
 - Burnup (GWd/MTHM)
 - Decay time, i.e., the time interval between the end of reactor use and emplacement in the repository
- Environmental concentrations of H₂, O₂, CO₃²⁻, and Fe²⁺
- Temperature

The model calculates a dose rate (J/kg) at the fuel surface induced by the radiation field. The dose rate is a function of the fuel properties and relationships presented in Radulescu (2011). The code runs from time zero to 100,000 years with a uniform log₁₀ temporal discretization.

Outputs are calculated at each time step. They include:

- Concentrations of UO₂²⁺, UO₂(CO₃)₂²⁻, UO₂(aq), U(IV)(s), U(VI)(s), H₂, O₂, H₂O₂, CO₃²⁻, and Fe²⁺ for each cell in the 1D domain

- Corrosion layer thickness [L], calculated from the cumulative amount of uranium precipitation (50% porosity assumed)
- UO_2 fuel degradation rate [$\text{M/L}^2/\text{T}$], calculated from the amount of fuel that reacts during the time step

Assumptions and Limitations

The FMD process model relies on several assumptions. They include:

- Fuel degradation in this model is exclusively due to the release of UO_2 from the UO_2 fuel matrix.
- Radiolysis effects are dominated by α -dose. This assumption breaks down for out-of-reactor times less than 30 years prior to contact with environmental water (Buck et al. 2013).
- α -particles are assumed to have a penetration distance of 35 μm and a constant energy of 5.3 MeV over that distance (Jerden et al. 2012).
- H_2O_2 (generated by radiolysis) and O_2 (introduced at the environmental boundary and generated by the decomposition of H_2O_2) are the only oxidants in the system.
- The corrosion layer partially reduces the active surface area of the fuel, partially blocks α -particles from generating radiolytic oxidants, and slows diffusion of reactants in the pores of the corrosion layer (Buck et al. 2013).
- Cladding is not included in the model.
- Fuel surface degradation is uniform.

The ranges of input values over which the FMD model may be applied have not been fully explored. The ranges shown in Tab. B-1 are theoretically valid. The ranges of concentrations of the environmental species are taken from documented example applications of the FMD process model or from personal communication from Jim Jerden (pers. email to Mariner, May 7, 2020 and June 16, 2020). Nearly all reactions in the FMD process model are kinetically controlled; therefore, combinations of aqueous species concentrations far from equilibrium are possible.

Tab. B-1 Theoretical FMD process model input ranges

| Parameter | Minimum | Maximum |
|--|-----------------------------------|----------------------------------|
| Initial Temp. (K) | 298 | 473 |
| Burnup (GWd/MTHM) | 20 | 90 |
| Environmental CO_3^{2-} (mol/m ³) | 10^{-4} (10^{-7} mol/liter) | 10^2 (10^{-1} mol/liter) |
| Environmental O_2 (mol/m ³) | 10^{-7} (10^{-10} mol/liter) | 10^0 (10^{-3} mol/liter) |
| Environmental Fe^{2+} (mol/m ³) | 10^{-3} (10^{-6} mol/liter) | 10^{-2} (10^{-5} mol/liter) |
| Environmental H_2 (mol/m ³) | 10^{-7} (10^{-10} mol/liter) | 10^0 (10^{-3} mol/liter) |

Coupling Requirements

To couple the FMD process model with PFLOTRAN, a “coupled” FMD process model was coded in Fortran. At each time step, PFLOTRAN calls the coupled FMD process model to obtain a new degradation rate. Coupling requires PFLOTRAN to keep track of the 1D chemical profiles across the domain from the previous time step for each breached WP. Other inputs include temperature, time, and several environmental concentrations in the boundary cell.

The specific requirements for implementation of the FMD process model in GDSA Framework are:

Inputs. The coupled FMD model must read the following for each WF of each breached WP at each time step:

- From the coupled FMD process model
 - Fuel burnup (GWd/MTHM)
 - Fuel decay time (time between the end of reactor use and repository emplacement)
 - Fuel specific surface area
 - Environmental concentrations of H_2 , O_2 , CO_3^{2-} , and SO_4^{2-} near the fuel surface
 - A future option will be to obtain these concentrations from PFLOTRAN when PFLOTRAN is used to simulate reactive transport of these species.
- From PFLOTRAN
 - Current time of simulation
 - Time step length
 - Temperature near the fuel surface
 - Concentrations of $UO_2(s)$, $UO_3(s)$, $UO_4(s)$, H_2O_2 , UO_2^{2+} , UCO_3^{2-} , UO_2 , CO_3^{2-} , O_2 , Fe^{2+} , and H_2 in each cell of the 1D process model domain between the fuel surface and the bulk water from the previous time step

Calculations. With these inputs, the coupled FMD process model must calculate at each time step:

- Dose rate for the time step as a function of burnup, decay time, and time of simulation
- Concentrations of $UO_2(s)$, $UO_3(s)$, $UO_4(s)$, H_2O_2 , UO_2^{2+} , UCO_3^{2-} , UO_2 , CO_3^{2-} , O_2 , Fe^{2+} , and H_2 in each cell of the 1D domain at the end of the time step

Outputs. The coupled FMD process model must return to PFLOTRAN at each time step:

- Fuel degradation rate ($kg\ m^{-2}\ s^{-1}$)
- Concentrations of $UO_2(s)$, $UO_3(s)$, $UO_4(s)$, H_2O_2 , UO_2^{2+} , UCO_3^{2-} , UO_2 , CO_3^{2-} , O_2 , Fe^{2+} , and H_2 in each cell of the 1D domain so that PFLOTRAN can store them for the next time step

FMD SURROGATE MODELS

Three types of surrogate models were developed for the FMD process model – a polynomial regression surrogate, an ANN surrogate, and a k-Nearest Neighbors regressor (kNNr) surrogate (Mariner et al.

2019). Of the two active learners (ANN and polynomial), the ANN surrogate performed better, so development of the polynomial surrogate was discontinued in FY 2020.

The section describes the development of the ANN and kNNr surrogates in standalone mode. Generation of training data for these surrogates is described in the Source Data section. The development and accuracy of the standalone kNNr and ANN surrogates are described in K Nearest Neighbors Regression (kNNr) and Artificial Neural Network (ANN) sections, respectively. The relative accuracy of these surrogates is compared in the Summary Comparison section.

Source Data

In FY 2021, the version 2 standalone MATLAB FMD model was used to generate a more focused set of training and testing data for the surrogate models. There were four main differences over the FMD source data used in Mariner et al. (2020). The first difference was the generation of a much larger training data set. The second and third differences involved “down sampling” the data in two dimensions to create a better balanced and more size efficient training model for the kNNr surrogate. The fourth change involved limiting the range of application and testing so that the training set extends beyond the edges of the application to reduce potential errors at the edges of the application range. The following sections describe the process for generating FMD source data and the four improvements.

Standalone MATLAB FMD Data Generation

As indicated in Fig. B-3, larger datasets reduce the distance to nearest neighbors in kNNr. To produce a larger dataset in FY 2021, the number of simulations was increased from 50,000 to 400,000. The standalone MATLAB version of the FMD model from 2018 (version 2) was used to generate the source data. The temporal discretization in each FMD execution (referred to as a “simulation” or “run”) consists of 101 logarithmically spaced (base 10) points from 0 to 10^5 years.

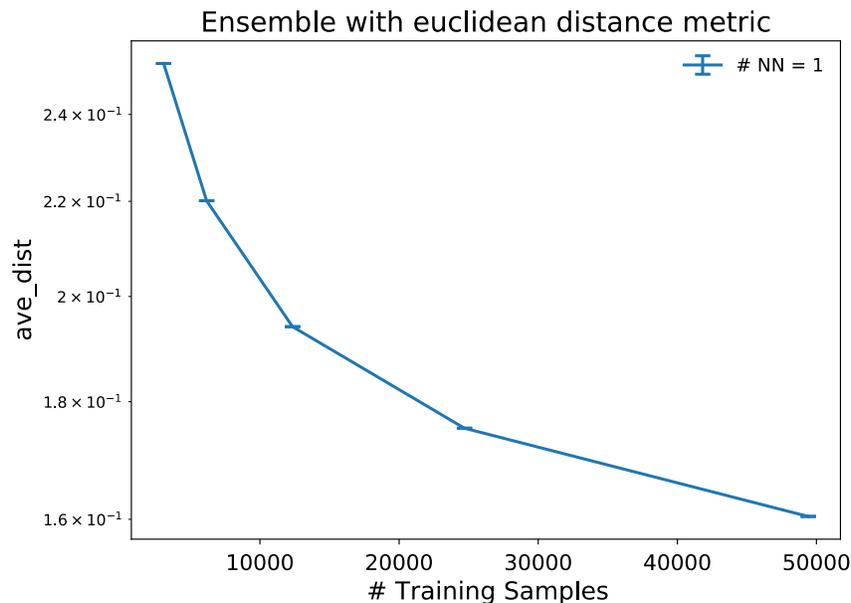


Fig. B-3 Average distance to the nearest neighbor in kNNr as affected by the number of training samples

For the new simulations, the ranges of initial temperature, burnup, and the environmental concentrations of CO_3^{2-} , O_2 , Fe^{2+} , and H_2 were reduced. For certain areas of the FY 2020 parameter space, the FMD MATLAB model has trouble converging, leads to unphysical corrosion layer thicknesses, or leads to stagnating runs, as observed in Mariner et al. (2020). Consequently, input variable ranges for the training data were changed from those shown in Tab. B-2 to those in Tab. B-3.

Tab. B-2

FMD process model input parameter ranges for surrogate training and testing FY 2020

| Parameter | Minimum | Maximum | Distribution |
|--|-----------------------------------|----------------------------------|--------------|
| Initial Temp. (K) | 298 | 393 | Uniform |
| Burnup (GWd/MTHM) | 20 | 70 | Uniform |
| Environmental CO_3^{2-} (mol/m ³) | 10^{-4} (10^{-7} mol/liter) | 10^{-1} (10^{-4} mol/liter) | Log-uniform |
| Environmental O_2 (mol/m ³) | 10^{-7} (10^{-10} mol/liter) | 10^{-3} (10^{-6} mol/liter) | Log-uniform |
| Environmental Fe^{2+} (mol/m ³) | 10^{-3} (10^{-6} mol/liter) | 10^{-2} (10^{-5} mol/liter) | Log-uniform |
| Environmental H_2 (mol/m ³) | 10^{-7} (10^{-10} mol/liter) | 10^{-1} (10^{-4} mol/liter) | Log-uniform |

Tab. B-3 FMD process model input parameter ranges for surrogate training FY 2021

| Parameter | Minimum | Maximum | Distribution |
|--|-----------------------------------|--|--------------|
| Initial Temp. (K) | 300 | 400 | Uniform |
| Burnup (GWd/MTHM) | 40 | 65 | Uniform |
| Environmental CO_3^{2-} (mol/m ³) | 10^{-3} (10^{-6} mol/liter) | 2×10^{-2} (2×10^{-5} mol/liter) | Log-uniform |
| Environmental O_2 (mol/m ³) | 10^{-7} (10^{-10} mol/liter) | 10^{-5} (10^{-8} mol/liter) | Log-uniform |
| Environmental Fe^{2+} (mol/m ³) | 10^{-3} (10^{-6} mol/liter) | 10^{-2} (10^{-5} mol/liter) | Log-uniform |
| Environmental H_2 (mol/m ³) | 10^{-5} (10^{-8} mol/liter) | 2×10^{-2} (2×10^{-5} mol/liter) | Log-uniform |

Training data was generated using a Latin Hypercube Sampling (LHS) design. LHS is a stratified sampling technique that generates “well-spaced” samples; it typically gives lower variance statistical estimators than plain Monte Carlo sampling (Helton and Davis 2003). The initial temperature (Kelvin) is a prescribed function of time (years) that exponentially decays from an initial value to 298 K:

$$T(t) = \begin{cases} T_0 - 110e^{-40t^{-0.5}}, & T > 298, \\ 298, & T \leq 298. \end{cases} \quad (\text{B-1})$$

After running the simulations, a filter was applied to remove runs that either took too long to converge, had corrosion layer thicknesses greater than 5 cm, or that stagnated, as discussed in Mariner et al. (2020).

Down Sampling on Time Trajectories

Fig. B-4 shows an example of how kNNr can result in biased predictions. If data are clustered in the sample space and all k nearest neighbors come from one cluster, then the neighboring clusters have no effect on the prediction. In this Fig. B-4, the clusters are due to high resolution along one dimension of the sample space relative to other dimensions. In the case of the FMD process model training data, the clusters are due to finely discretized time series outputs.

To improve the balance of data densities across the multidimensional sample space, the time trajectories produced by process model simulations were down sampled to retain only a randomly selected portion of the time points. Down selection was set at 9 of the 101 data points per run. This down selection ratio not only improved kNNr predictions, it resulted in a smaller training data set for the kNNr model which improved performance. After down sampling, the training data set resulted in about 2.3 million training points.

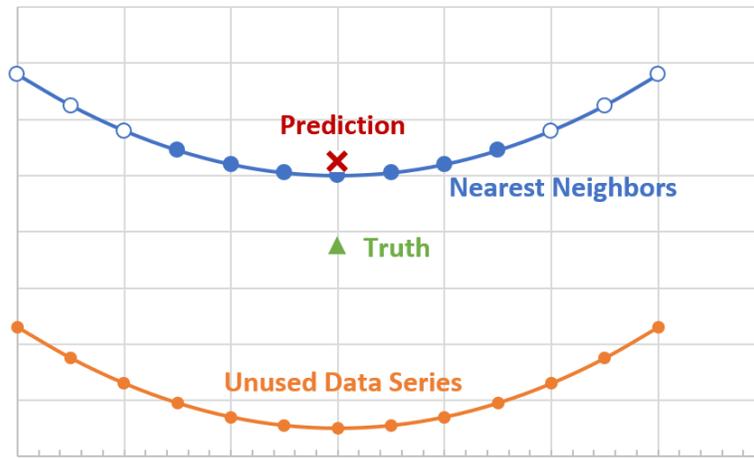


Fig. B-4 An example of kNNr error when the sample space contains clusters of training data.

Sample Space Dimension Reduction

A correlation analysis showed that there were significant differences in the contributions of input parameters to the output parameter (Fig. B-5). Iron and oxygen showed essentially zero correlation with the output parameter. By excluding these inputs, kNNr prediction accuracy improved significantly as discussed in more detail in the section on kNNr hyper parameter tuning. In addition, dropping these two inputs reduced the size of the kNNr training model and increased the efficiency of the surrogate model.

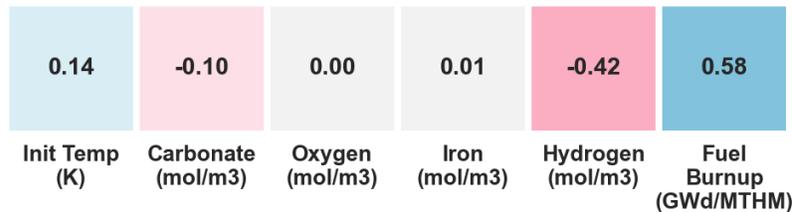


Fig. B-5 Correlation of input parameters with the predicted uranium oxide surface flux.

Training beyond the Ranges of Application

Examination of the dataset where prediction accuracy suffered in FY 2020 led to the realization that kNNr interpolation was not performing well at the edges of the training space. To avoid potential edge effects, the ranges of application and testing were narrowed in FY 2021 to the ranges shown in Tab. B-4. The test set was not down sampled and resulted in about 1.2 million data points.

Tab. B-4 FMD process model input parameter ranges for surrogate application and testing FY 2021

| Parameter | Minimum | Maximum | Distribution |
|---|--|---|--------------|
| Initial Temp. (K) | 310 | 390 | Uniform |
| Burnup (GWd/MTHM) | 45 | 60 | Uniform |
| Environmental CO ₃ ²⁻ (mol/m ³) | 2 × 10 ⁻³ (2 × 10 ⁻⁶ mol/liter) | 10 ⁻² (10 ⁻⁵ mol/liter) | Log-uniform |
| Environmental O ₂ (mol/m ³) | 2 × 10 ⁻⁷ (2 × 10 ⁻¹⁰ mol/liter) | 5 × 10 ⁻⁶ (5 × 10 ⁻⁹ mol/liter) | Log-uniform |
| Environmental Fe ²⁺ (mol/m ³) | 2 × 10 ⁻³ (2 × 10 ⁻⁶ mol/liter) | 5 × 10 ⁻³ (5 × 10 ⁻⁶ mol/liter) | Log-uniform |
| Environmental H ₂ (mol/m ³) | 2 × 10 ⁻⁵ (2 × 10 ⁻⁸ mol/liter) | 10 ⁻² (10 ⁻⁵ mol/liter) | Log-uniform |

Difference Between kNNr and ANN Source Data

The ANN surrogate employed all but one of the source data changes used by the kNNr surrogate. ANN retained all six input parameter dimensions. This allowed using code from FY 2020 and likely made little performance difference because the ANN training should reject such data during the training phase by assigning low weights to those features. The kNNr surrogate model relies upon a type of lookup table comprising a carefully selected subset of the FMD generated training data; its performance is tied very closely to data quality and coverage. The ANN surrogate involves the development of an internal, functional representation of the source data during training that enables the ANN to smooth over some source data issues.

Error Metrics

Three error metrics were calculated to assess the surrogate model predictions of the test data: mean-squared error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). The formulas for these are given below for reference in terms of model predictions h_i and truth values y_i :

$$MSE = \frac{1}{N} \sum_{i=1}^N (h_i - y_i)^2 \quad (\text{B-2})$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |h_i - y_i| \quad (\text{B-3})$$

$$MAPE = \frac{100}{N} \sum_{i=1}^N \left| \frac{h_i - y_i}{y_i} \right| \quad (\text{B-4})$$

K Nearest Neighbors Regression (kNNr)

The accuracy of the k-nearest neighbors regressor (kNNr) as a surrogate model was improved this year to levels favorably comparable to the ANN regressor. The following sections give an overview of kNNr development, describes the process for tuning kNNr parameters, and examines the performance of the improved kNNr surrogate.

This effort builds upon prior work from Mariner et al. (2020), developing, tuning, and coupling a kNNr based surrogate model with the state-of-the-art, massively parallel Fortran code called PFLOTRAN in support of nuclear waste repository analysis. The feasibility of achieving significant speedup with limited loss of accuracy was previously demonstrated by precomputing a large number of states and implementing fast surrogate models using k-nearest neighbors and ANNs regressors (Mariner et al. 2019).

By identifying the nearest points in the precomputed data and interpolating and returning the corresponding uranium oxide degradation rates, the kNNr model functions as a type of high speed look up table and FMD surrogate model. It is coupled to the massively parallel subsurface flow and reactive transport code, PFLOTRAN, so that PFLOTRAN can repeatedly interrogate the kNNr surrogate UO₂ values. To effectively balance the need for both speed and accuracy, the kNNr model hyperparameters, such as the amount of training data and the number of nearest neighbors, were tuned to provide acceptable predictions overall, though in certain regions of the sample space, more improvement is needed. The priority for FY2021 was to improve overall accuracy with a stretch goal of characterizing where the surrogates can be trusted and where to proceed with caution.

Overview of kNNr

The kNNr (Ben-David and Shalev-Shwartz 2014) is a supervised, non-parametric machine learning method that, unlike polynomial regression or neural networks, does not re-express the data in any way in order to make predictions. In contrast to the ANN approach, which employs active learners, the k-Nearest Neighbors regressor is a lazy learner that tabulates data points inside of a domain X with labels y to the end of using those values for predictions. This makes the kNNr highly interpretable, as no intermediate hypothesis selection process on the parameters is undertaken. Instead, the label for a point within the domain but not in the “table” is obtained as an average of the labels of the *k* nearest neighbors of this new point, where $k \geq 1$ is fixed (Fig. B-6).

The definition of nearest depends on the metric function used, though a typical choice is the Minkowski metric $(\sum_{i=1}^d |x_i - y_i|^p)^{\frac{1}{p}}$, with $p \geq 1$. The case of $p = 2$ is the popular Euclidean metric, whereas $p = 1$ gives the Manhattan distance. The tabulation of data points can be implemented with a matrix representing entries in a table. However, this is less efficient than modern tabulation methods like the k-d Tree or the Ball Tree (Pedregosa, Varoquaux et al. 2011). The actual calculation of the predicted value need not be a uniform average. An inverse of the distance to each neighbor may be used to determine how influential that neighbor is in the final calculation of the weighted average.

One of the attractive features of kNNr is that it makes predictions based on local information only, and therefore does not require global smoothness over the input space. On the other hand, the approach requires a sufficiently dense table to get good predictive accuracy, and the cost of table look-ups increases as the table density increases.

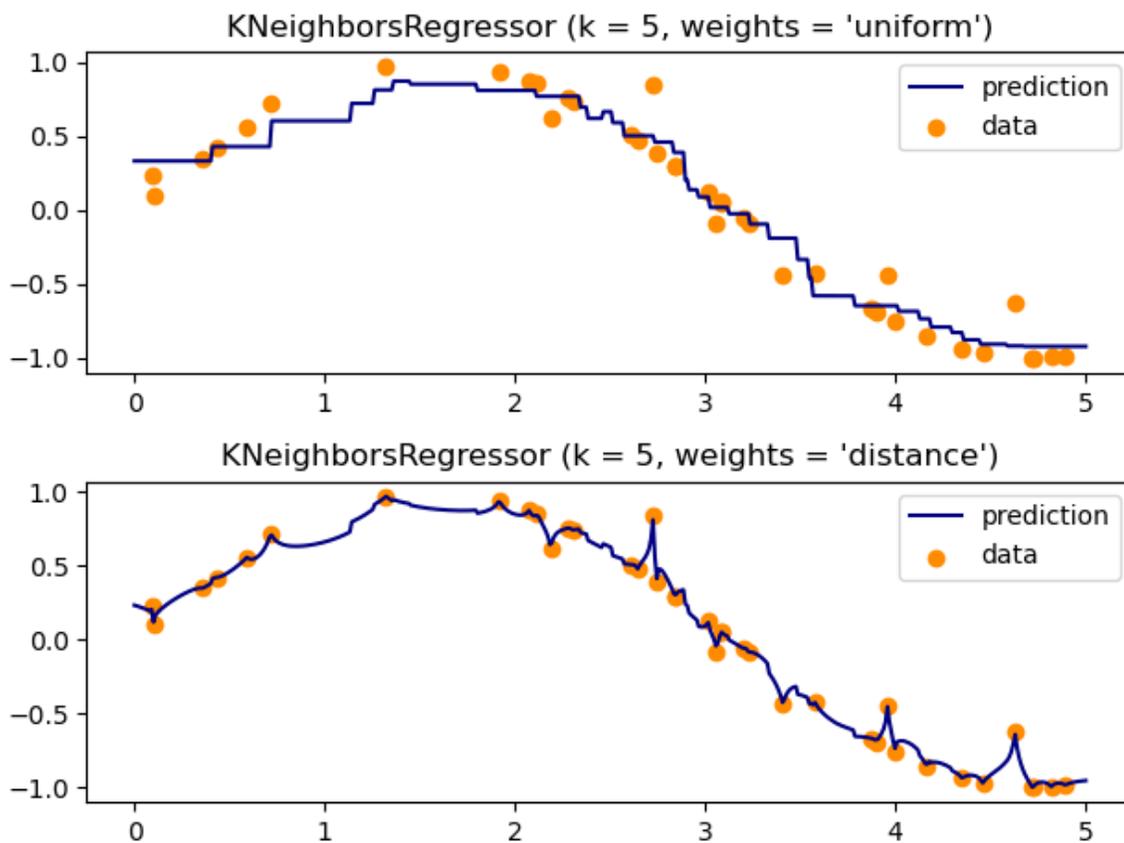


Fig. B-6 While k-nearest neighbors is often thought of as a classification algorithm, it is also used extensively for regression. Predictions of “missing” values are derived by interpolating with a selected number of nearest neighboring data points. In the one-dimensional examples above, the top chart shows the predictions obtained by averaging five nearest neighbors to interpolate missing values. In the lower plot, the interpolation weights the five neighbors based on the inverse of their distance to desired, missing value. In this case, parameter selections appear to overfit the predictions to the data. This image is from the scikit-learn documentation at https://scikit-learn.org/stable/auto_examples/neighbors/plot_regression.html

Development and Hyper-Parameter Tuning

The same Python and Fortran kNNr tools were used as in Mariner et al. (2020), consisting of three versions of kNNr: a Python kNNr version for rapid prototyping of features and parameter tuning, a standalone version of the production Fortran kNNr for feature improvements and verification testing, and the production Fortran version of kNNr coupled to PFLOTRAN.

Rapid Prototyping in Python with Scikit-Learn kNNr

The rapid prototyping in Python performed in FY 2020 continued using a corresponding kNNr model implemented with the Python Scikit-Learn module (Pedregosa et al. 2011). Scikit-Learn offers many, built-in features not available in Fortran for tuning hyperparameters and calculating prediction accuracy. The standalone Fortran and Python versions were used in tandem offline with the Fortran version for informing computational and timing costs and the corresponding Python-based kNNr used for broader hyperparameter tuning. Verification testing in Mariner et al. (2020), confirmed that the Fortran and Python versions were functioning compatibly.

KDTREE 2 for Production Fortran kNNr

As detailed in Mariner et al. (2020), the open source module KDTREE 2 (Kennel 2004) is a Fortran version built for speed and with a limited but adequate set of kNNr features for testing the feasibility of the approach. Developed at the University of California, San Diego Institute for Nonlinear Science, KDTREE 2 is a Fortran95 module written with “considerable care... in the implementation of the search methods, resulting in substantially higher computational efficiency (up to an order of magnitude faster) than the ... previous version.” Specifically, “the k-d data structure and search algorithms are the generalization of classical binary search trees to higher dimensional spaces, so that one may locate near neighbors to an example vector in $O(\log N)$ time instead of the brute-force $O(N)$ time, with N being the size of the data base” (Kennel 2004). The intended advantages of the code align well with the FMD use cases where the database is large, higher dimensional, and requires larger numbers of nearest neighbors. The author states that “the improvements are the most potent in the more difficult and slowest cases” which is essentially the FMD case. The algorithm is based on Cormen’s popular text book (Cormen et al. 2009). The KDTREE 2 Github repository is forked by several developers with the master archived in 2009 here: <https://github.com/jmhodges/kdtree2/tree/master/src-f90>.

The KDTREE 2 library as available on GitHub does miss some key features. For example, it did not come with any interpolation capabilities. As such, as detailed in Mariner et al. (2020), a function to provide inverse distance weighted averaging between the retrieved nearest neighbors was added. The code was also expanded to accept HDF5 binary data files.

Data Conditioning, Hyperparameter Tuning and Timing Tests

Several, interdependent data conditioning, hyperparameter tuning and evaluation criteria were used. Similar to Mariner et al. (2020), an inverse-distance weighted interpolation between the retrieved nearest neighbors was used to make a prediction from the tabulated data for a query vector. Also, based on experiments performed in Mariner et al. (2020), the k-d Tree algorithm was used for data partitioning, along with the Euclidean distance metric, which are both available in the Fortran KDTREE 2 code. Further, following Mariner et al. (2020), the same data conditioning process was used. Runs with unphysical corrosion layer thicknesses were removed, as well as duplicates in the feature set. All features and the Quantity of Interest (QoI) were transformed into \log_{10} space, with both the prediction and interpolation done in the transformed space to get better accuracy.

One key difference with Mariner et al. (2020) is the nature and size of the training data used here. As discussed in Mariner et al. (2020), retaining all time samples from a standalone FMDM process model run led to a situation where the data density was much higher in the Temperature and dose rate dimensions than in the environmental concentration space. As such, very often most of the nearest neighbors selected for a query point were at the same environmental concentrations instead leading to low prediction accuracy as there was not enough variability in the concentrations of the selected nearest neighbors for a good interpolation.

For the kNNr metaparameter tuning, a second LHS-sampled data set of 20,000 standalone fuel cask runs was generated. As shown in Tab. B-4, this validation data set was generated over a slightly smaller parameter space in an effort to reduce extrapolation at the edges of the training data set. However, these runs were not subsampled in time. A third LHS-sampled data set of 20,000 standalone fuel cask runs over the same parameter range as the validation set was used for testing data.

Hyperparameter Tuning: Nearness Metric, Number of Neighbors and Training Dataset Size

To determine the best combination of hyperparameters and to average over the noise induced by selecting random subsets of data, ensembles of tests were run varying the neighbor nearness metric, number of nearest neighbors, and amount of training data. Balancing the potential accuracy of using large amounts

of training data versus the speed advantages of smaller sets is part of the hyperparameter tuning. Since the size of the training data set impacts the accuracy of different numbers of nearest neighbors, they were addressed simultaneously. To balance the probabilistic effects of choosing training subsets at random (one might test really well and another poorly), ensembles were used and their statistics calculated. The predictions were then compared with the true values of the validation data QoI or target values to compute pointwise, run averaged, and ensemble errors along with their statistics.

A first set of tuning runs was done with all 6 features in the training data set: temperature, environmental concentrations of CO_3^{2-} , O_2 , Fe^{2+} , and H_2 , and the dose rate at the fuel surface. The results of these runs are shown in Fig. B-7 and Fig. B-8. They show that both MSE and MAPE error metrics decrease with more samples in the training data. The errors for a training data set with 9 time samples per run are slightly better than those with 3 samples per run. For the case with 9 samples per run, the lowest errors are achieved with 2.25 million training samples and 55 nearest neighbors, with an MSE error of about 2.9×10^{-6} and a MAPE error of about 75%.

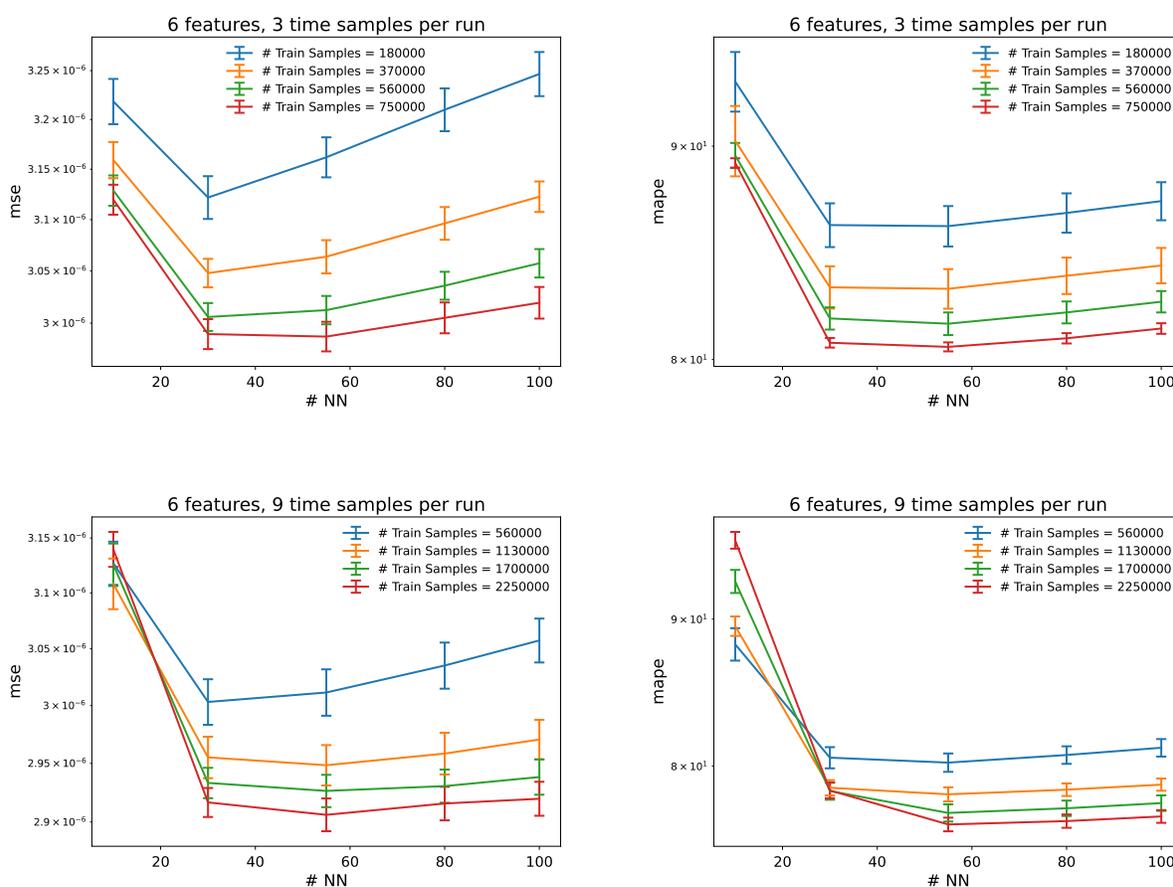


Fig. B-7 MSE and MAPE errors plotted as a function of the number of nearest neighbors for ensembles of ten randomly selected training datasets varying in size, using 6 features in the training data.

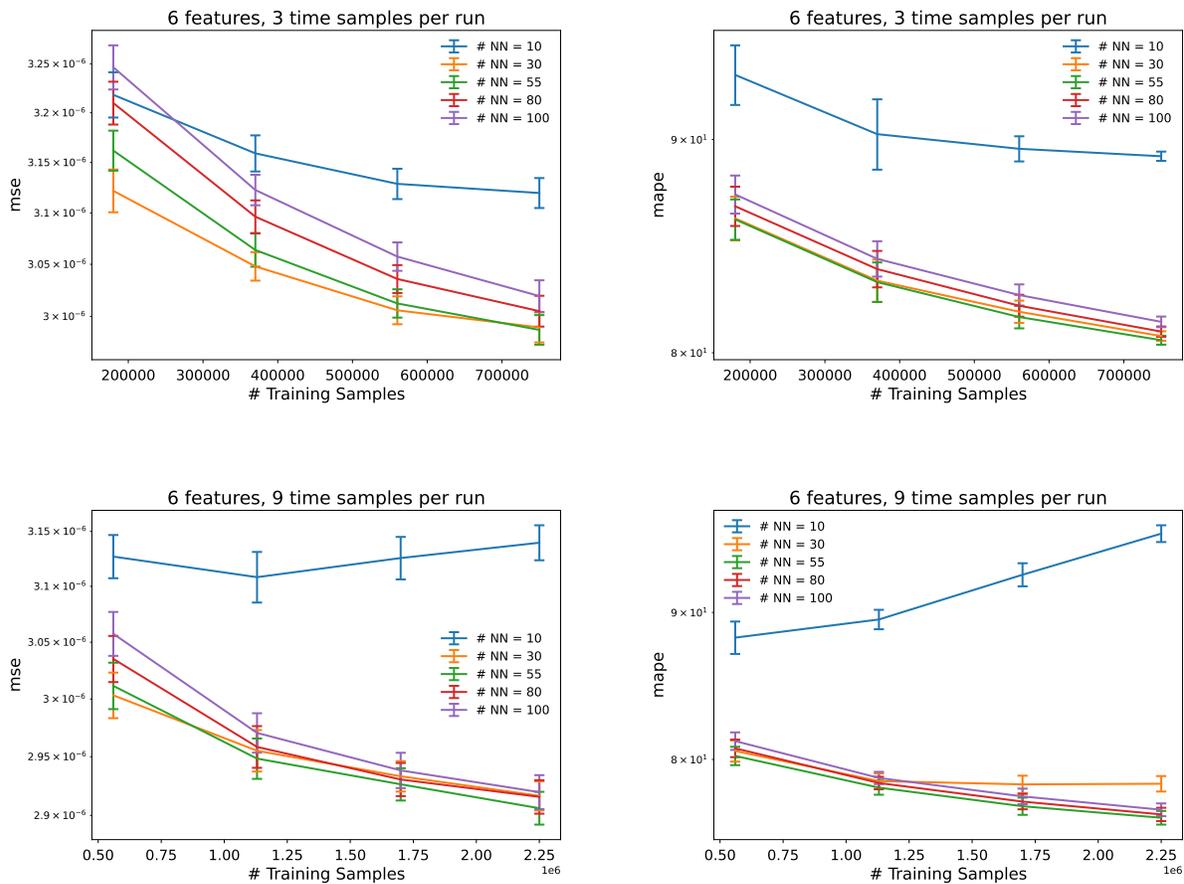


Fig. B-8 MSE and MAPE errors plotted as a function of the number of training samples used for ensembles of ten randomly selected training datasets, using 6 features in the training data.

Fig. B-9 and Fig. B-10 show the results of this same analysis for the scenario where only the 4 most relevant features in the training data set are see temperature, environmental concentrations of CO_3^{2-} , and H_2 , and the dose rate at the fuel surface. Similar trends are observed as with the previous results, but the error metrics look substantially better (especially the MAPE errors). In general, adding more training data results in lower errors, as long as enough nearest neighbors are used in the interpolation. The MSE error is lowest at about 80 nearest neighbors, whereas the MAPE does best at 30 nearest neighbors. For this case with 4 features, optimal results are also seen for the data sets with 9 time samples per run.

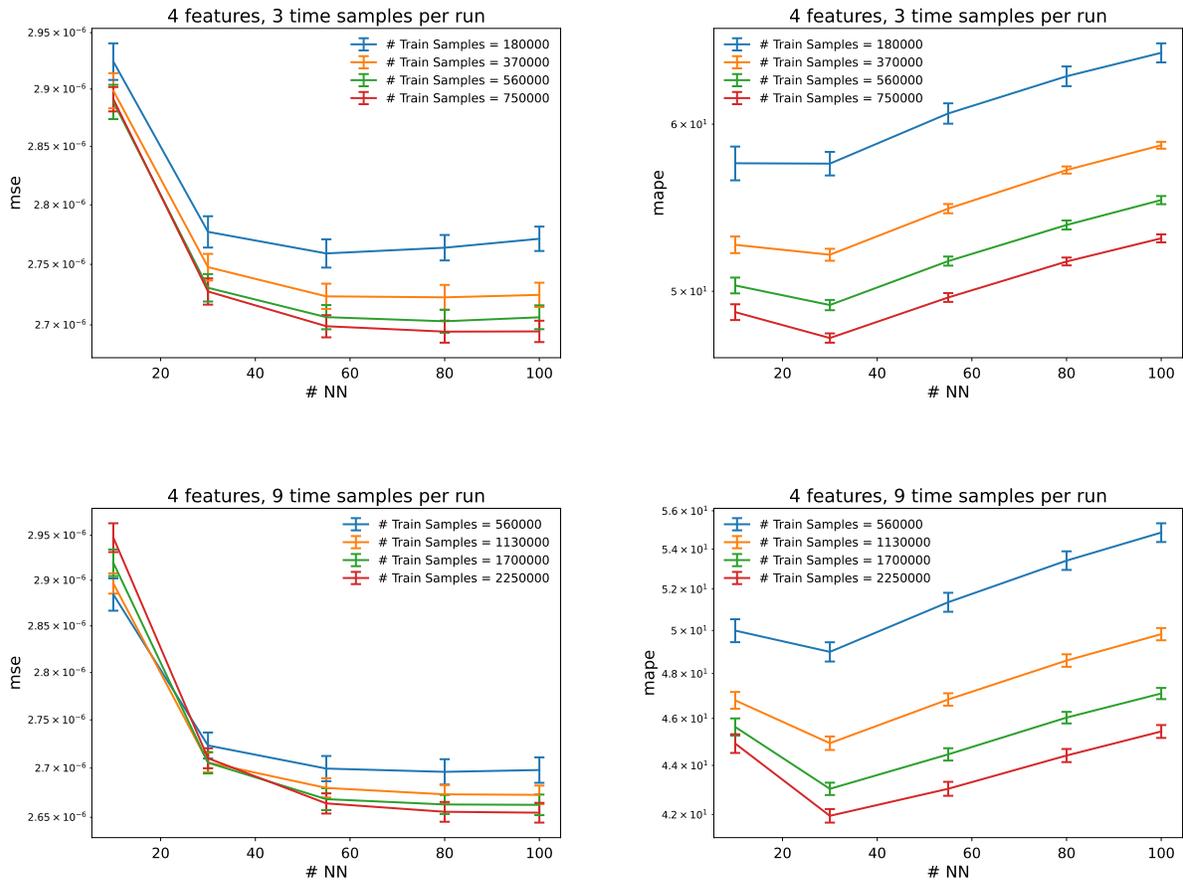


Fig. B-9 MSE and MAPE errors plotted as a function of the number of nearest neighbors for ensembles of ten randomly selected training datasets varying in size, using 4 features in the training data.

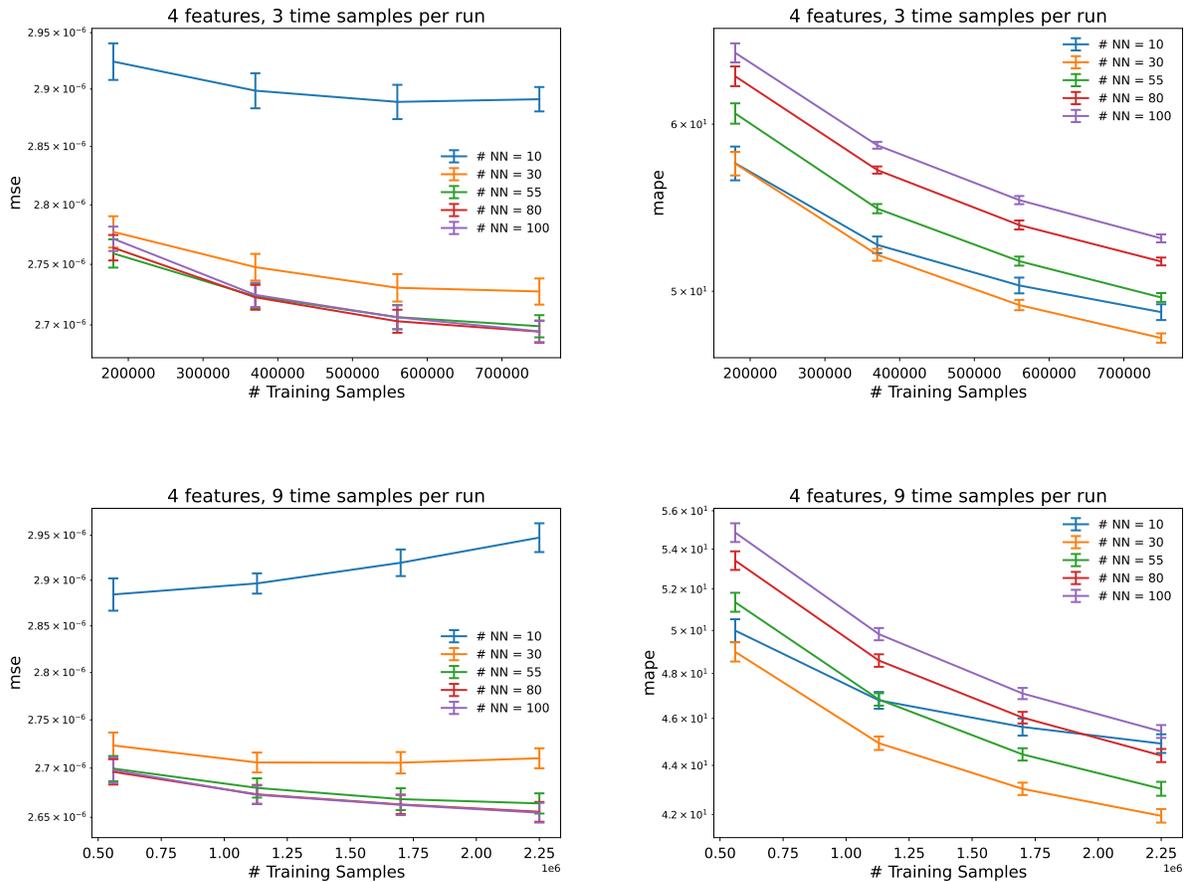


Fig. B-10 MSE and MAPE errors plotted as a function of the number of training samples used for ensembles of ten randomly selected training datasets, using 4 features in the training data.

This is confirmed in Fig. B-11, which compares the MSE for the largest data sets for each degree of subsampling for both the case with 4 and 6 features. Optimal results are obtained for the training data set that has 9 random time samples per run, for a total of 2.25 million samples, using 80 nearest neighbors, resulting in a MSE of 2.65×10^{-6} . The optimal MAPE error is about 42% for this same data set but using only 30 nearest neighbors.

As the MSE error is the most relevant for the accuracy of the coupled reservoir simulations, the configuration with 9 samples per time run, using the 4 features (temperature, environmental concentrations of CO_3^{2-} , and H_2 , and the dose rate at the fuel surface), using 80 nearest neighbors is chosen to move forward.

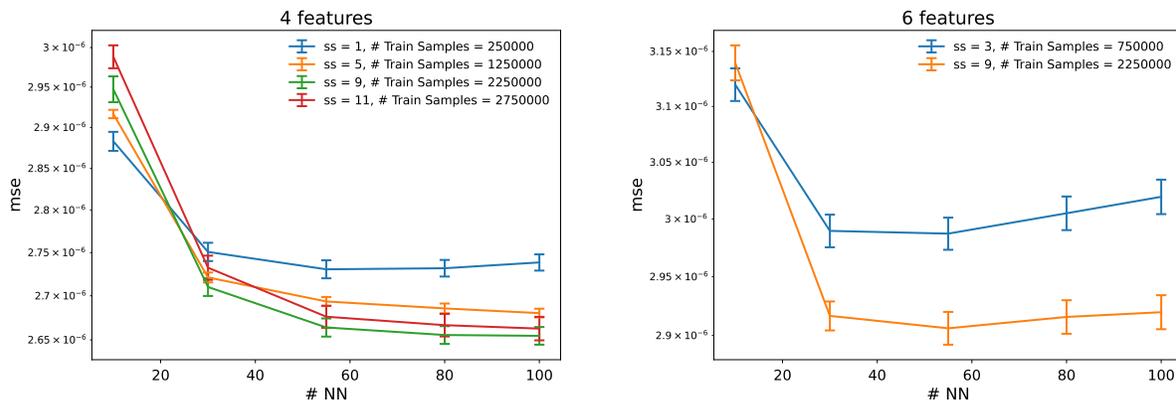


Fig. B-11 MSE errors plotted as a function of the number of nearest neighbors used for ensembles of ten randomly selected training datasets varying in size, using 4 (left) or 6 (right) features in the training data.

Accuracy

Since the MSE error does better as more training data are made available, all subsequent kNNr predictions were made with the full set of samples in the data set with 9 randomly selected time samples per run. Overall, this data set had 2.28 million samples, and we accordingly reference the optimal configuration as kNNr 2.28M/80NN going forward. To gain insights into the errors on the test set, one hundred random trajectories predicted with this kNNr 2.28M/80NN configuration are compared to the true trajectories in Fig. B-12.

While some predictions appear to be working reasonably well, there are zones where performance is poor. Looking more closely at the best, worst, and median predictions (on a per-run averaged basis) in Fig. B-13 reveals runs with considerable errors. Since kNNr is essentially a look up table, for these cases the algorithm must be operating in areas where it lacks the density of training data required.

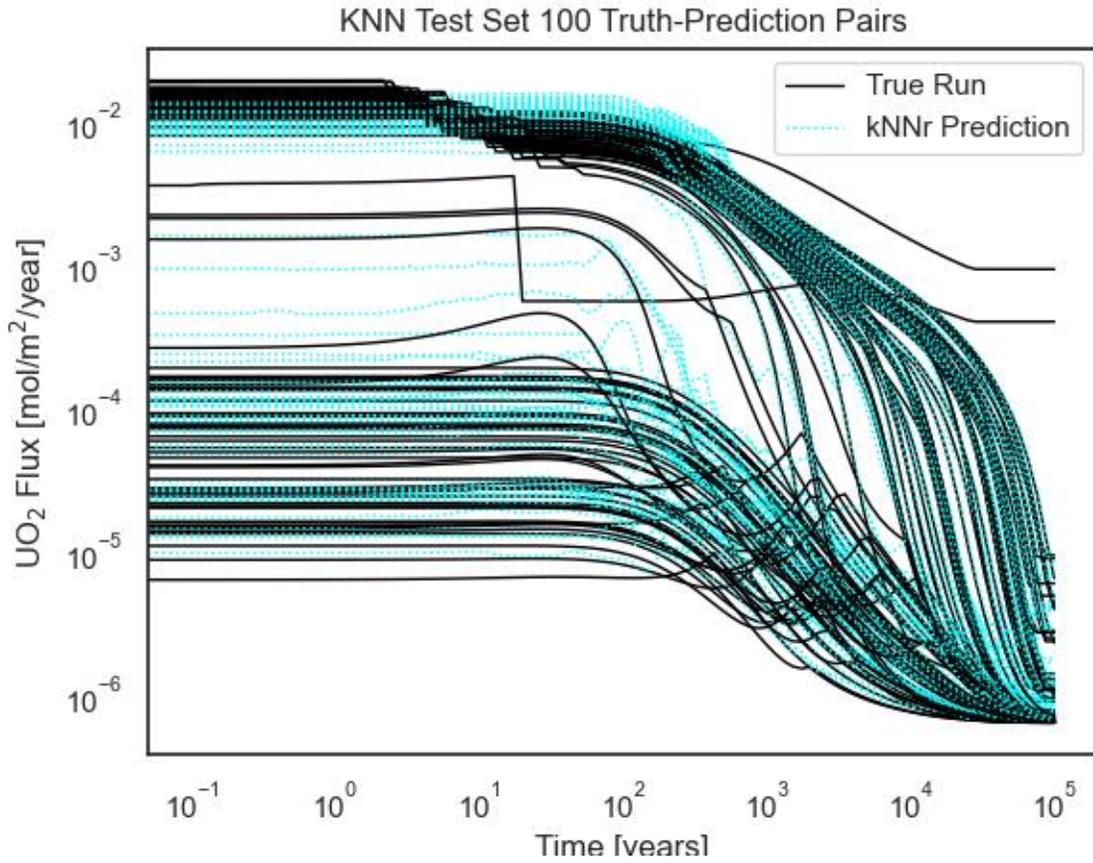


Fig. B-12 One hundred randomly selected kNNr truth prediction pairs shows a range of FMD trace shapes (True Runs) and that kNNr Prediction accuracy varies considerably.

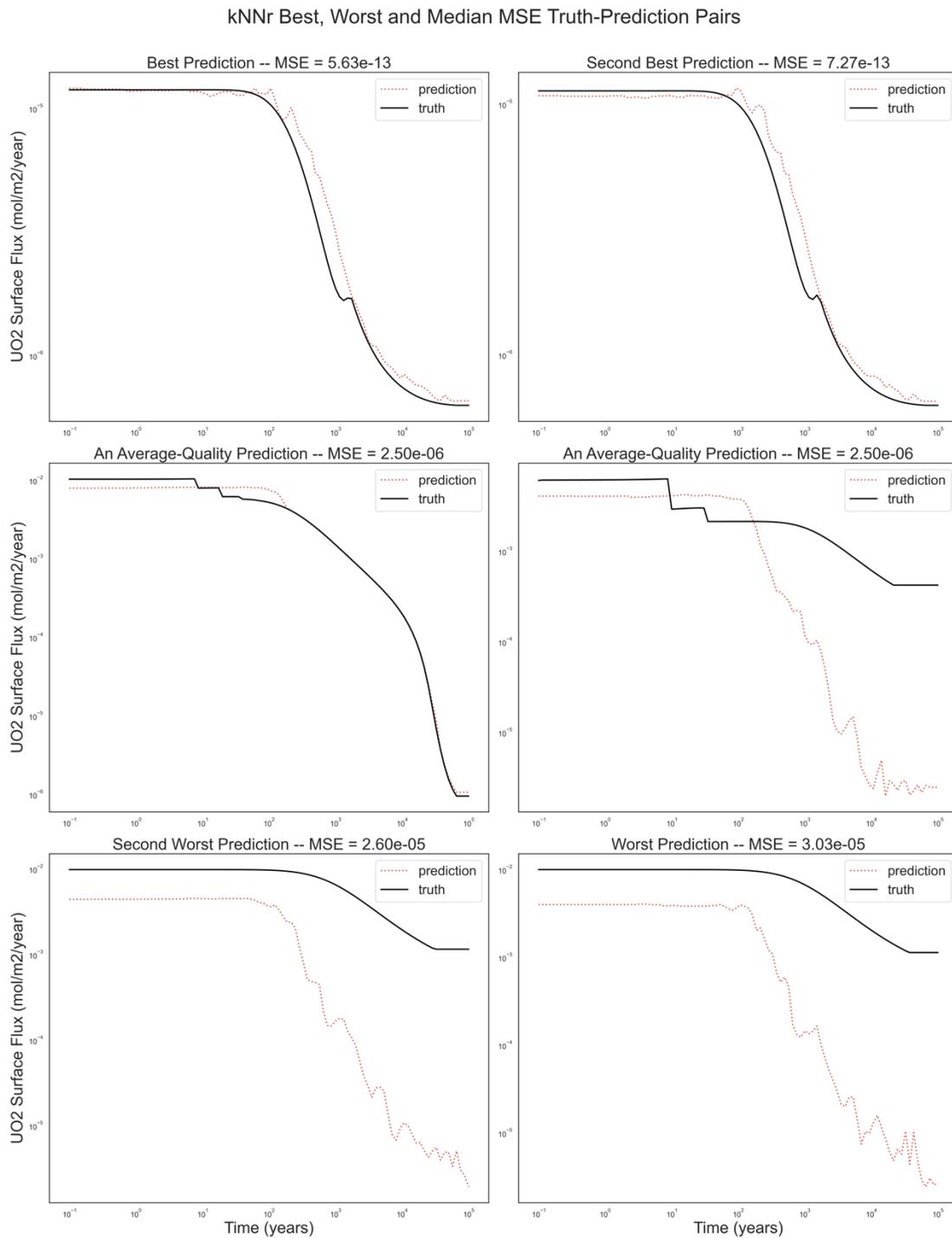


Fig. B-13 Plots of the best, median and worst kNNr runs by prediction mean squared error (MSE)

To get a better feel for the error distributions, Fig. B-14 plots the histograms of the run-averaged MSE, MAPE, and MAE errors. When binning the errors themselves, as in the left column, the histogram shows a peak near zero, and then a long tail towards larger errors. By binning the 10-based log of the errors, we get a better insight into the behavior of the tails. Both the MSE and MAE error metrics show large secondary peaks in the error distributions. Future work will focus on these peaks and other outliers to see where the table does not perform well, and adjust the sampling strategy, distance metric, or data conditioning accordingly.

k-NNr Test Dataset Error Metrics

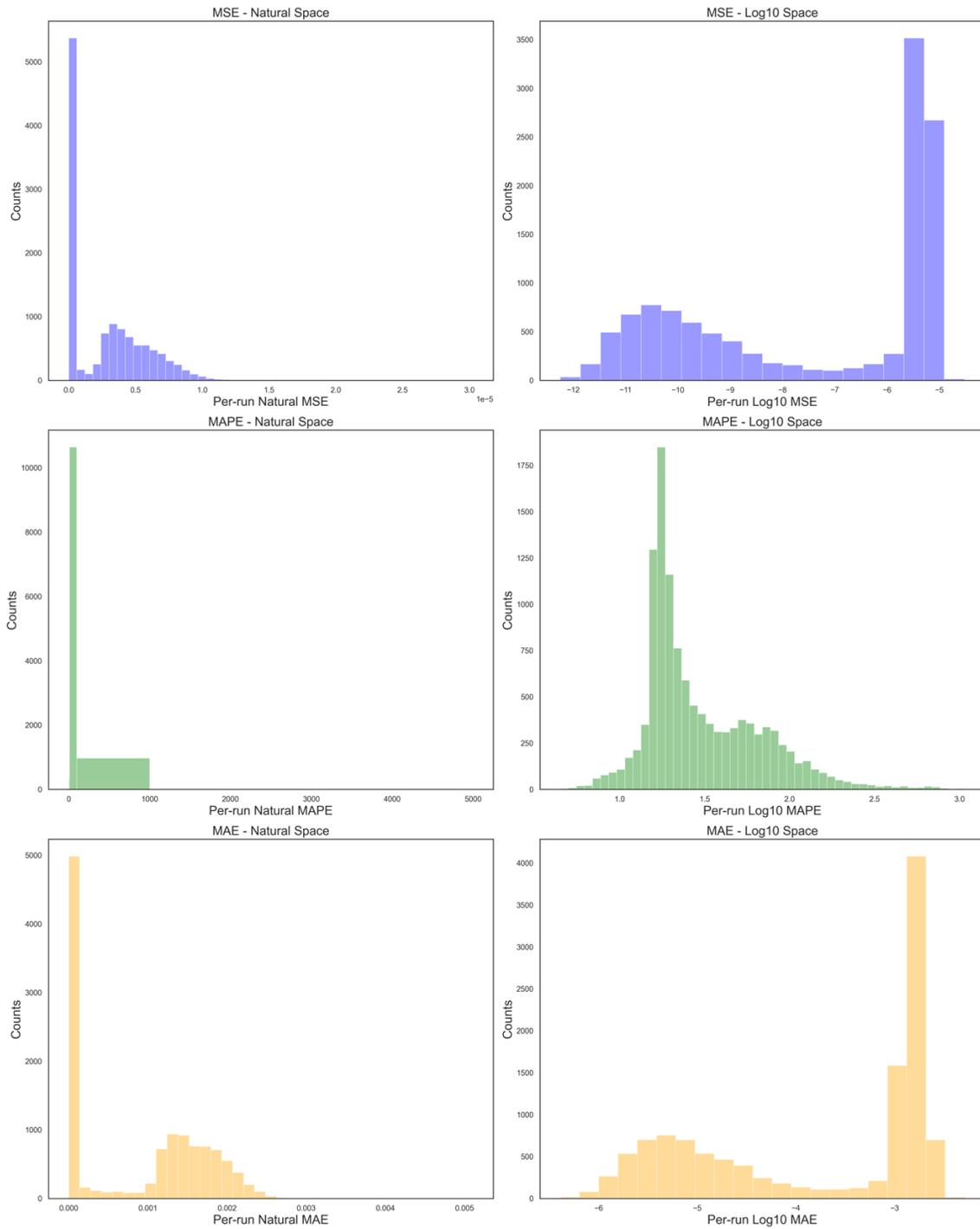


Fig. B-14 Histogram plots of the kNNr prediction errors continues to show a bi-modal error distribution indicating there is a large set of predictions that are much less accurate.

Artificial Neural Network (ANN)

Neural network models are commonly employed by the machine learning community for regression and classification problems. They can be described as intricate networks of “artificial neurons” that are essentially weighted combinations of (usually simple) nonlinear functions. One motivation for the development of neural networks (Rasmussen and Williams 2006; Pedregosa et al. 2011; Ben-David and Shalev-Shwartz 2014) was to create a regression approach for complex functions that avoids the combinatorial growth of the feature space that occurs in polynomial regression models.

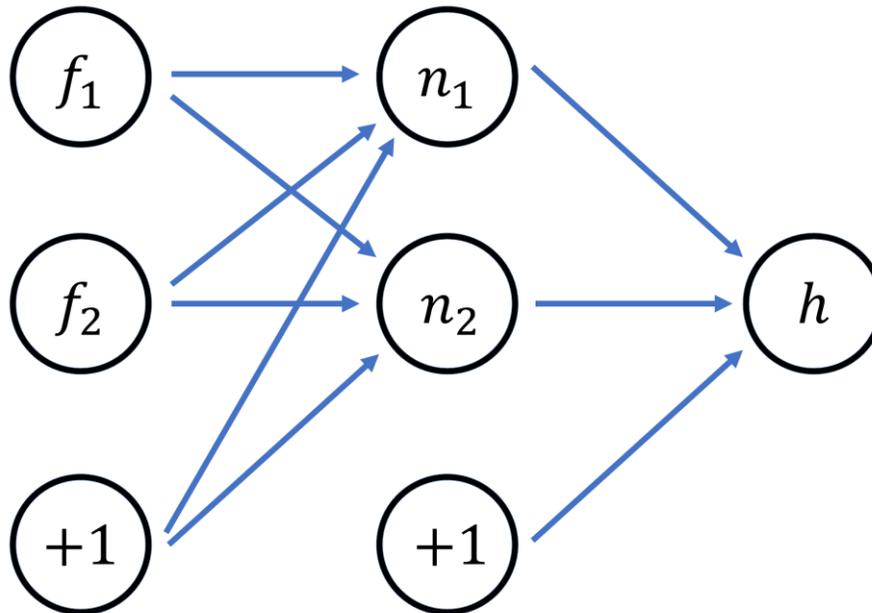


Fig. B-15 A schematic of a single layer feed-forward neural network with 2 input features and 2 neurons in the hidden layer

Fig. B-15 contains a depiction of a single layer feed-forward neural network with two features and a single output. It is called a single layer network because there is one “hidden” layer of neurons between the input and output layers, and the term feed-forward reflects that all the “weights”, the directed connections between neurons, are pointing from the input layer to the output layer. The +1 nodes denote the “bias” or “offset” terms that are independent of the features f_i and hidden layer neurons. The “depth” of a network is defined as the number of hidden layers. So-called “deep” neural networks contain at least two hidden layers (and often more), while networks that contain only a few layers are described as “shallow.”

The inputs to a neuron are scaled by their corresponding weights w_{ij} (which also includes the bias term for convenience), summed, and then fed into a nonlinear “activation function.” The index i denotes the layer in the network and j the node or bias term in a given layer. In this work we use the popular rectified linear unit (ReLU) activation function, which is zero for an input less than zero and equal to the input otherwise. The output of each neuron in the final hidden layer is weighted and summed at the output node to produce the model prediction. In regression (as opposed to classification) neural networks there is typically no activation function applied at the output node.

The process of training a neural network involves minimizing a loss function that depends on the weights and biases w_{ij} (blue arrows in Fig. B-15).

Development & Training

Similar to Mariner et al. (2020) we utilize a mean absolute error (MAE) loss function for the training:

$$J(w_{ij}) = \frac{1}{N} \sum_{k=1}^N |h(w_{ij}, x_k) - y_k| \quad (\text{B-5})$$

where N denotes the number of data points in the training set and $h(w_{ij}, x_k)$ is the neural network prediction for features x_k with corresponding QoI value y_k . All of the features and QoI were \log_{10} -transformed prior to training save for temperature, as they are strictly positive and vary over orders of magnitude. The features were also standardized to have zero mean and unit variance prior to training. MAE was chosen as the loss metric for training because it is less sensitive to outliers than mean squared error (MSE). We used the optimization algorithm RMSprop (Tieleman and Hinton 2012) to minimize (5) and terminated the algorithm after the improvement on the training set with additional iterations began to significantly diminish.

Based on convergence studies in Mariner et al. (2020) a network with two hidden layers and 64 nodes per layer was selected for the ANN surrogate model. This model contains 4673 parameters (weights and bias terms).

Accuracy

After determining the architecture of the network, the ANN was trained on the data set filtered from 400,000 runs and down sampled to 9 random samples per time trace. This is the same training data set used in the kNNr analysis. Predictions were computed on a test set of 20,000 standalone FMDM runs, also the same as was used in the kNNr analysis. One difference from the kNNr approach is that all 6 available features were used. The ANN is not expected to be negatively affected by additional features that may not be informative.

Fig. B-16 shows one hundred randomly selected surrogate prediction and test data traces from the FMD model, and Fig. B-17 displays the two best, worst, and average predictions with respect to MSE for a given simulation trace from the test set (i.e., “per-run” MSE). Fig. B-18 contains histograms of the per-run error metric. While most of the ANN predictions have a per-run error that is close to 0, there are some outliers.

One avenue for improvement of the ANN surrogate would be consideration of more complicated network architectures although there is no guarantee that this would be fruitful. Alternatively, diagnosing why the FMD model is stagnating, producing excess corrosion layer thickness, or not finishing in a timely manner, and addressing these issues would be helpful because it would remove the gaps caused by the filtering procedures in the training set which would likely improve prediction accuracy in those regions of feature space.

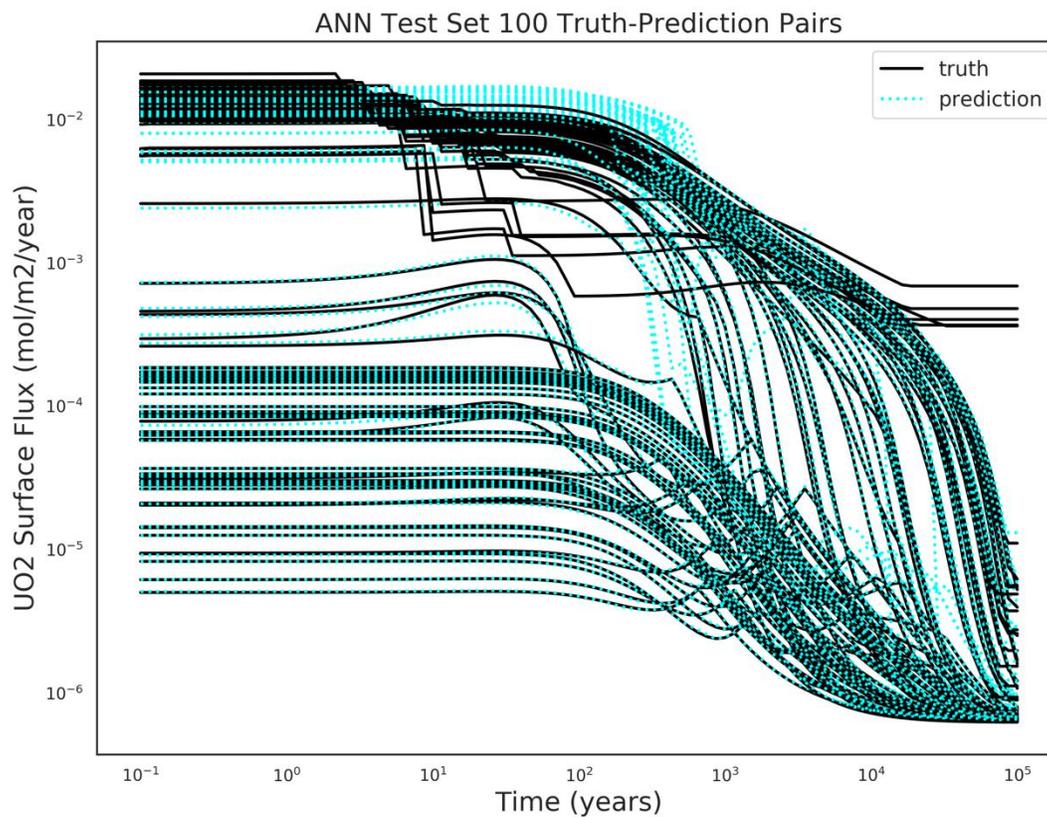


Fig. B-16 One hundred randomly-selected truth-prediction simulation trace pairs from the test set

ANN Individual Truth-Prediction Pairs

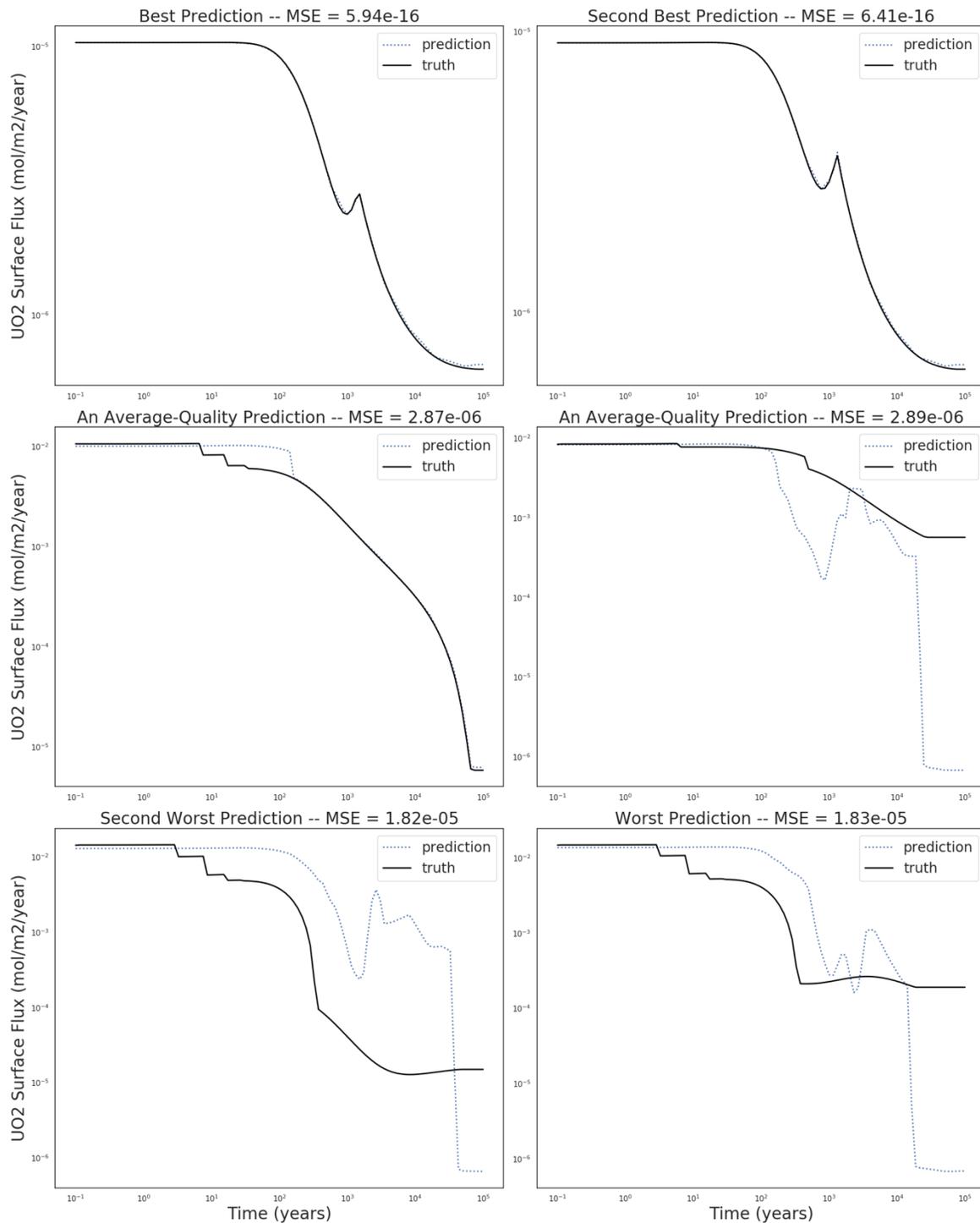


Fig. B-17 Individual truth-prediction pairs for the ANN surrogate

ANN Test Set Error Metrics

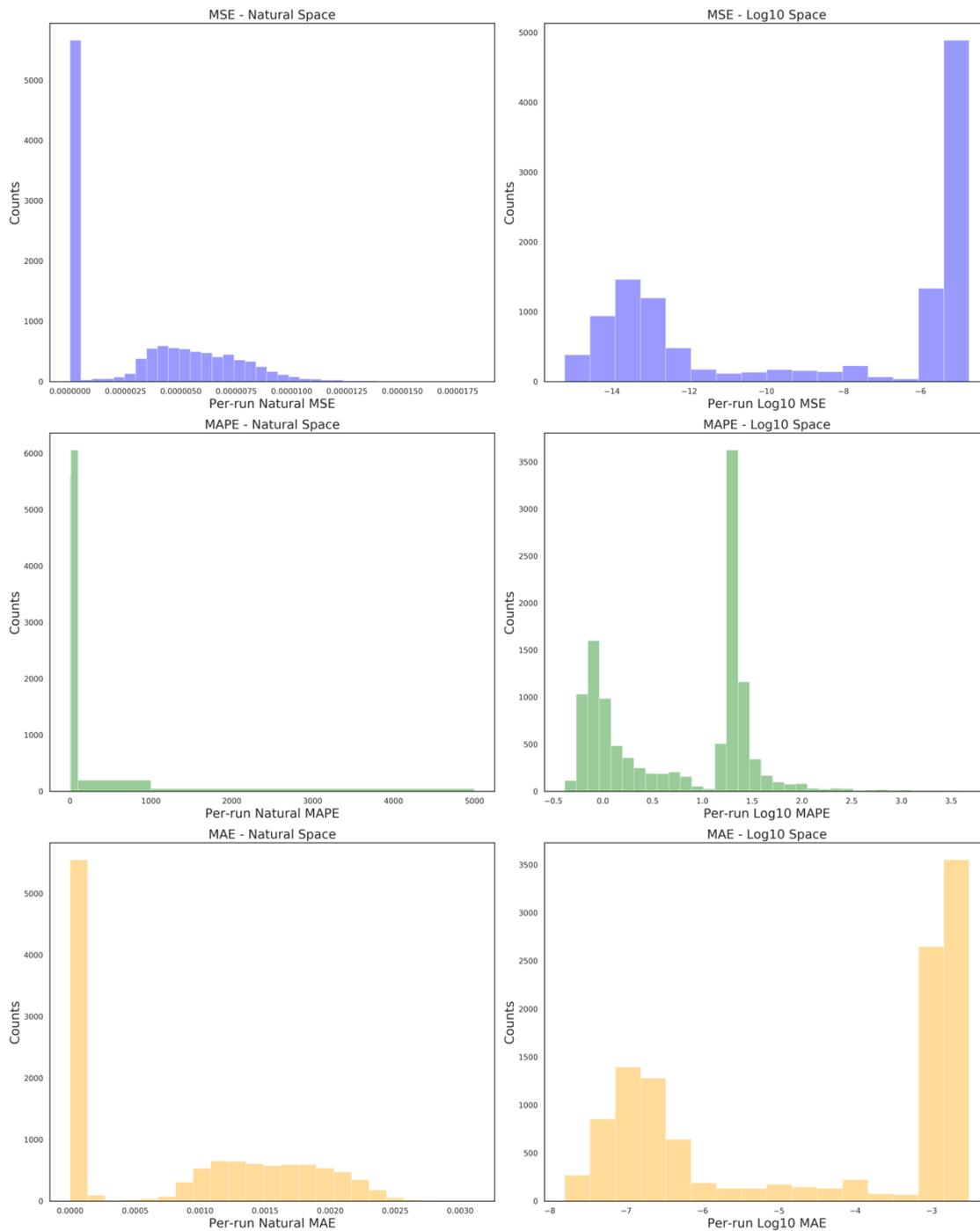


Fig. B-18 Distributions of the per-run (i.e., per FMD simulation trace) error metrics for the test set

Summary Comparison

Tab. B-5 compares the accuracy of ANN and kNNr surrogates on the test data set. For the data sets used here, the kNNr method is slightly more accurate in the MSE metric than the ANN method. The ANN surrogate shows better accuracy in the MAPE error. In the MAE metric, the ANN is slightly more accurate than kNNr. Thanks to the improved sampling approach, the MSE error in the optimal kNNr approach is less than half of what it was in Mariner et al. (2020).

Tab. B-5 Values of the error metrics computed on the test set (and training set for the ANN) for the UO₂ surface flux in the “natural” (i.e., non-log₁₀ transformed) space. Note that the error on the training data is zero for kNNr as we use inverse-distance weighted averaging so the table prediction is the same as the tabulated value at the query point.

| Surrogate | Train MSE (mol/m ² /yr) ² | Test MSE (mol/m ² /yr) ² | Train MAE (mol/m ² /yr) | Test MAE (mol/m ² /yr) | Train MAPE | Test MAPE |
|-----------------|--|---|--|--------------------------------------|---------------|--------------|
| ANN | 3.41e-6 | 3.08e-6 | 8.57e-4 | 8.26e-4 | 33.7% | 25.1% |
| kNNr 2.28M/80NN | N/A | 2.69e-6 | N/A | 8.92e-4 | N/A | 44.4% |

COUPLING FMD SURROGATES TO PFLOTTRAN

There are a few details that must be accounted for to achieve proper coupling between PFLOTTRAN and the surrogate models. First, the temperature input to the surrogate must be evaluated at the WP spatial location. The local environmental concentrations of CO₃²⁻, O₂, Fe²⁺, and H₂ must also be evaluated, but currently they remain constant over the simulation domain. Second, two unit conversions between PFLOTTRAN and those expected by the surrogate are necessary: mol/liter to mol/m³ for concentrations and degrees Celsius to degrees Kelvin for temperature. Lastly, all of the surrogate features except dose rate are directly available from PFLOTTRAN. Dose rate at the fuel surface is obtained by passing burnup, simulation time, and a decay time offset into a function that computes the dose rate at the fuel surface according to the formula in Radulescu (2011).

Unlike the actual FMD process model, there is no notion of a time step in the surrogate models. Both of surrogates are a functional representation (i.e., a mapping) between the feature space and model output. There is no history-dependence and none of the features are rate-like quantities. This is a particularly attractive property of these surrogate models given the restrictive time discretization (log₁₀ time increments) required by the FMD process model, which PFLOTTRAN must adhere to for the FMD model to work correctly. If a surrogate model is used instead then PFLOTTRAN is free to adaptively form the time discretization.

Details of the implementation of the ANN and kNNr surrogate models in PFLOTTRAN are presented in Mariner et al. (2020). Online documentation of implementation and user options can be found at https://doc-dev.pflotran.org/theory_guide/pm_waste_form.html and https://doc-dev.pflotran.org/user_guide/cards/gdsa/waste_form_general_card.html.

DEMONSTRATION

The coupled kNNr and ANN surrogate models are demonstrated in a full-scale shale repository reference case simulation. The purpose of these demonstrations is to show the effects of these models on the fuel

degradation rates over time for breached WPs and to assess the computer time needed for the surrogate calculations relative to flow and transport calculations.

The shale repository simulation includes (Sevougian et al. 2019):

- Natural and EBS in a shale formation
- 24-PWR and 37-PWR WPs
- Inventory of radionuclides and WFs used in the 2019 PA simulation model

Inputs

The model domain is 7215 m (L) × 2055 m (W) × 1200 m (H), long enough to implement observation points 5000 m down-gradient of the repository in the x-axis direction. The numerical grid contains 9.88 million cells, of which approximately 4.6 million are the finer cells in and around the repository. The repository is discretized into volumetric cells with size of 5 m (L) × 5 m (W) × 5 m (H) while most of other regions are discretized into 15 m (L) × 15 m (W) × 15 m (H) cells.

Initial pressure and temperature fields represent geothermal temperature and hydrostatic pressure gradients in a vertical direction (positive y-axis direction) by applying a liquid flux of 0 m/s and an energy flux of 60 mW/m² to the bottom and atmospheric pressure and 10°C of temperature at the top boundary. Pressure at the top decreases from west to east (positive x-axis direction) with a head gradient of -0.0013 m/m. Initially, unsaturated conditions (0.7 of gas phase saturation) are applied in the disposal drifts, hallways, and shafts, and concentrations of all radionuclides in all cells are 10⁻²¹ mol/liter.

The symmetry boundary condition is applied to the south face (positive zx-face) whereas other boundaries have constant pressure and temperature. Radionuclide concentrations are 10⁻²¹ mol/liter at the upstream boundary and zero at the outflow boundaries. The simulation runs to 10⁶ years.

This field-scale model simulates WP degradation, WF dissolution, equilibrium-controlled radionuclide sorption, precipitation/dissolution, radioactive decay and ingrowth in all phases, and multi-physical coupling of heat transfer, fluid flow and radionuclide transport.

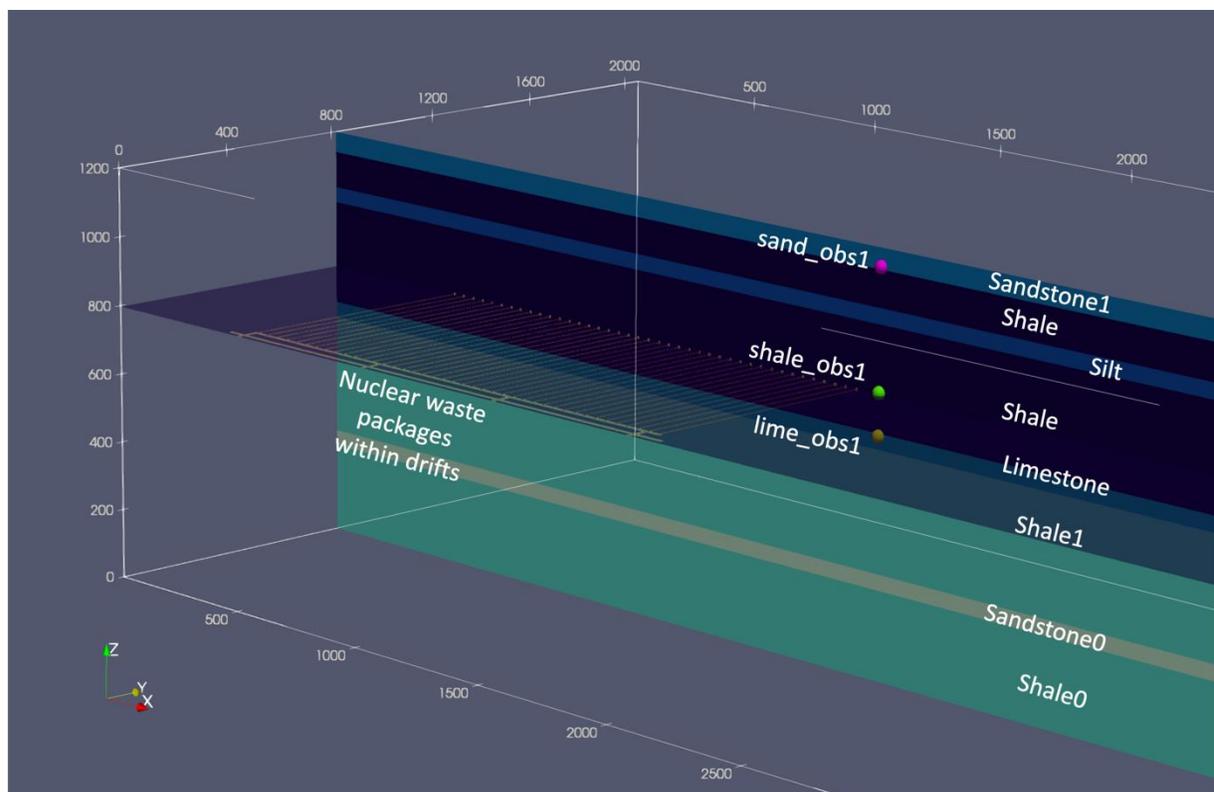


Fig. B-19 Model domain of the shale repository system

Formation Properties

The nuclear waste repository is located at depth of 405 m within the shale formation. The stratigraphic units of the layered system are shown in Fig. B-19 and the material properties are given in Tab..

In this modeling study, the disturbed rock zone (DRZ) is defined manually by assuming 10 times larger permeability than that of the host rock to consider the changes in rock properties adjacent to the EBS.

Waste Packages

The shale reference case implements 1575 24-PWR and 1000 37-PWR WPs consisting of stainless-steel canisters and overpacks. Note that only half of the WPs are modeled due to the symmetry-domain setting. The radionuclide inventory at the time of emplacement and heat of decay as function of time are calculated via decay and ingrowth from the 5-year out of reactor (OoR) radionuclide inventories (Sevougian et al. 2019). Calculated values for 24-PWR WPs assume an initial enrichment of 3.72 wt% ^{235}U , 40 GWd/MTHM burn-up, and 100-year OoR surface storage prior to deep geologic disposal. Calculated values for 37-PWR WPs assume an initial enrichment of 4.73 wt% ^{235}U , 60 GWd/MTHM burn-up, and 150-year OoR storage.

The WPs are modeled as cuboids ($1.67 \times 1.67 \times 5 \text{ m}^3$) to resolve gridding limits. The porosity of WPs is set to 0.5, which is equal to the fraction of void space. Its permeability is set to $1 \times 10^{-16} \text{ m}^2$, several orders of magnitude higher than that of the surrounding materials.

The temperature-dependent degradation rate per year with a truncated log normal distribution (a mean of -4.5, a standard deviation of 0.5, and an upper truncation of -3.5 in log units) is implemented to calculate normalized remaining canister thickness (fractional thickness) at each time step.

Tab. B-6 Material properties for the shale repository system model

| | Host Shale | Upper sandstone (sandstone1) | Silty shale (silt) | Limestone (Limestone) | Lower shale (shale1) | Lower sandstone (sandstone0) | Bottom shale (shale0) |
|---|------------------------|------------------------------|--------------------|-----------------------|----------------------|------------------------------|-----------------------|
| Location [m] (Domain top at 1200 m and bottom at 0 m) | | 1140-1200 | 990-1035 | 645-690 | 300-645 | 255-300 | 0-255 |
| Permeability [m ²] | 1e-19 | 1e-13 | 1e-17 | 1e-14 | 1e-20 | 1e-13 | 1e-20 |
| Porosity [-] | 0.2 | 0.2 | 0.2 | 0.1 | 0.1 | 0.2 | 0.1 |
| Density[kg/m ³] | 2700 | 2700 | 2700 | 2700 | 2700 | 2700 | 2700 |
| Heat capacity | 830 | 830 | 830 | 830 | 830 | 830 | 830 |
| Saturated thermal conductivity [W/m/K] | 1.2 | 3.1 | 1.4 | 2.6 | 1.2 | 3.1 | 1.2 |
| Dry thermal conductivity [W/m/K] | 0.6 | 1 | 0.8 | 1 | 0.6 | 1 | 0.6 |
| Liquid residual saturation | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| Gas residual saturation | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| Saturation function | Van Genuchten function | | | | | | |
| alpha [Pa ⁻¹] | 6.67e-7 | 1e-4 | 6.67e-7 | 1e-4 | 6.67e-7 | 1e-4 | 6.67e-7 |
| m [-] | 0.333 | 0.5 | 0.333 | 0.5 | 0.333 | 0.5 | 0.333 |

Bentonite Buffer

Compacted bentonite, filling access drifts and shafts, was engineered to have low permeability and high sorption capacity by mixing 70% bentonite and 30% quartz sand for 24-PWR disposal and adding 15% graphite for 37-PWR disposal.

Both buffer materials have a porosity of 0.35 and a permeability of 10^{-20} m². The bentonite/sand buffer has a saturated thermal conductivity of 1.5 W/m/K and a dry thermal conductivity of 0.6 W/m/K whereas the bentonite/graphite buffer has 3.0 and 1.9 W/m/K, respectively.

Geochemical Conditions

For the surrogate mechanism simulation, four environmental species are added to the chemical database and input file as follows:

Tab. B-7 Additional primary species for surrogate mechanism simulations

| | Ion size (a0) | Charge (Z) | Molar Weight (g/mol) | Free Conc. (mol/m ³) | Free Conc. (mol/liter) |
|-------------------------------|------------------|---------------|----------------------------|-------------------------------------|---------------------------|
| O ₂ (aq) | 3.0 | 0.0 | 31.9988 | 1e-6 | 1e-9 |
| H ₂ (aq) | 3.0 | 0.0 | 2.0159 | 1e-2 | 1e-5 |
| Fe ²⁺ | 6.0 | 2.0 | 55.8470 | 3e-3 | 3e-6 |
| CO ₃ ²⁻ | 4.5 | -2.0 | 60.0092 | 1.4e-2 | 1.4e-5 |

The updated database is named as “ufd-decay_ann.dat” for PFLOTRAN simulations.

Surrogate Mechanism Inputs

Tab. B-8 Surrogate mechanism settings for the shale repository system model

| PFLOTRAN Keyword | 24-PWR | 37-PWR |
|-----------------------|----------------------------|----------------------------|
| MATRIX_DENSITY | 10970 [kg/m ³] | 10970 [kg/m ³] |
| SPECIFIC_SURFACE_AREA | 0.001 [m ² /g] | 0.001 [m ² /g] |
| BURNUP | 40 [GWd/MTHM] | 60 [GWd/MTHM] |
| DECAY_TIME | 100 [year] | 100 [year] |

Results

Degradation Rate

The original shale reference case uses a simple fractional degradation rate (FDR) model for fuel matrix degradation. FDR specifies a rate of 10⁻⁷ yr⁻¹ for the WF matrix in this simulation. This model and this 10⁻⁷ yr⁻¹ rate is used, for example, in the Swedish repository performance assessment model (SKB 2006, Table 10-3).

Fig. B-20 shows the time when each WP breaches in the simulation and the WF degradation rate history for each WP. The yellow and orange lines are for the 37-PWR WPs and the blue and green lines are for the cooler 24-PWR WPs. The top and bottom graphs plot the fuel matrix degradation rate in units of kg/s and mol/m²/yr, respectively. Because a constant specific surface area is assumed (0.001 m²/g), the bottom graph shows that as soon as the WP breaches, the fuel matrix degrades at a constant rate of approximately 3.7×10⁻⁷ mol/m²/yr for the remainder of the simulation. The middle graph tracks the remaining fuel matrix volume for each WP.

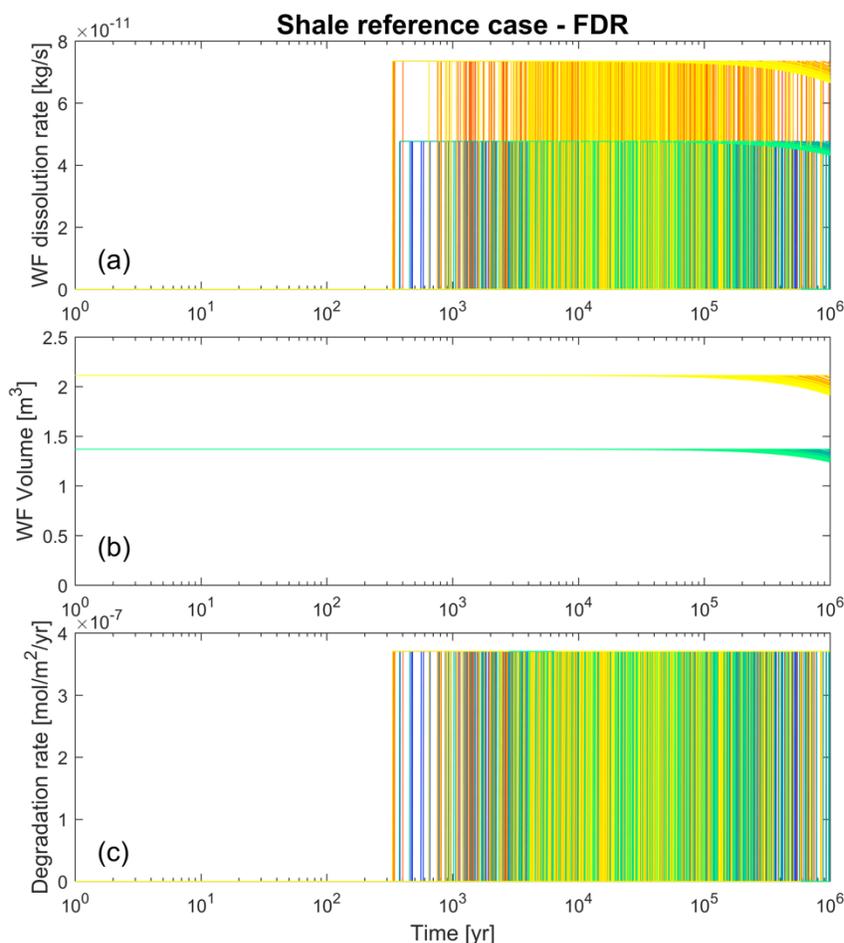


Fig. B-20 FDR model results: (a) fuel matrix degradation rate (kg/s), (b) volume of remaining waste form (m^3), and (c) specific degradation rate ($mol/m^2/yr$)

Fig. B-21 shows the same three graphs for the same shale reference case simulation, but instead of using the FDR fuel matrix degradation model, the ANN FMD surrogate model is used. The ANN surrogate calculates much higher fuel degradation rates than the FDR simulation, especially at early times. Because the surrogate emulates the FMD process model, the specific degradation rate ($mol/m^2/yr$) is high at early times when the temperature ($>100^\circ C$) and dose rates are high, and it decreases over time as temperature and dose rate decrease. These trends are expected and are a major advantage over the FDR model. Also, the rates for the 24-PWR WPs are generally lower than the rates for the 37-PWR WPs because of differences in temperature. ANN rates for both 24-PWR and 37-PWR converge after 100,000 years to approximately $6.5 \times 10^{-7} mol/m^2/yr$.

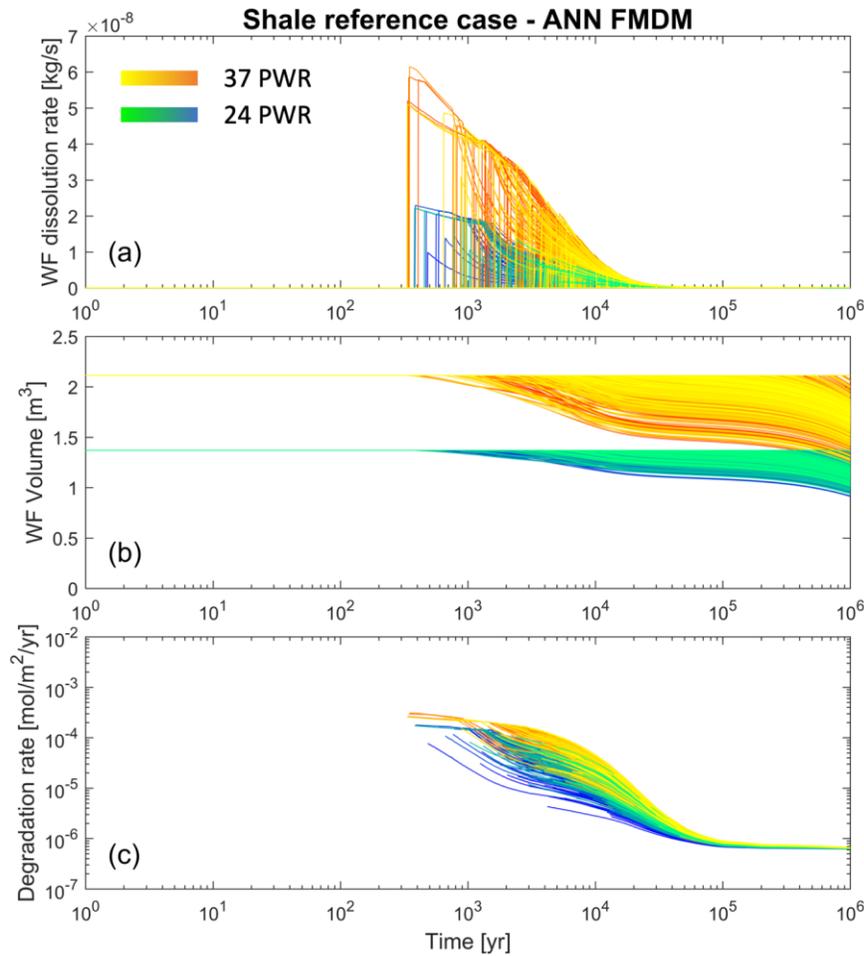


Fig. B-21 ANN FMD surrogate model results: (a) fuel matrix degradation rate (kg/s), (b) volume of remaining waste form (m^3), and (c) specific degradation rate ($mol/m^2/yr$)

Fig. B-22 shows the same plots for shale reference case when the kNNr surrogate model is used. The results show similar trends in degradation rates (high rates upon WP breach and then gradual decreases). The maximum value of the specific degradation rate ($\sim 2 \times 10^{-4} mol/m^2/yr$) is approximately the same as the maximum specific degradation rate predicted by the ANN surrogate. Also, the value over time converges to approximately the same value as observed in the ANN surrogate model.

The kNNr surrogate predicts lower degradation rates than the ANN surrogate over the period between approximately 2000 and 20,000 years. Further work is needed to determine the cause of this discrepancy.

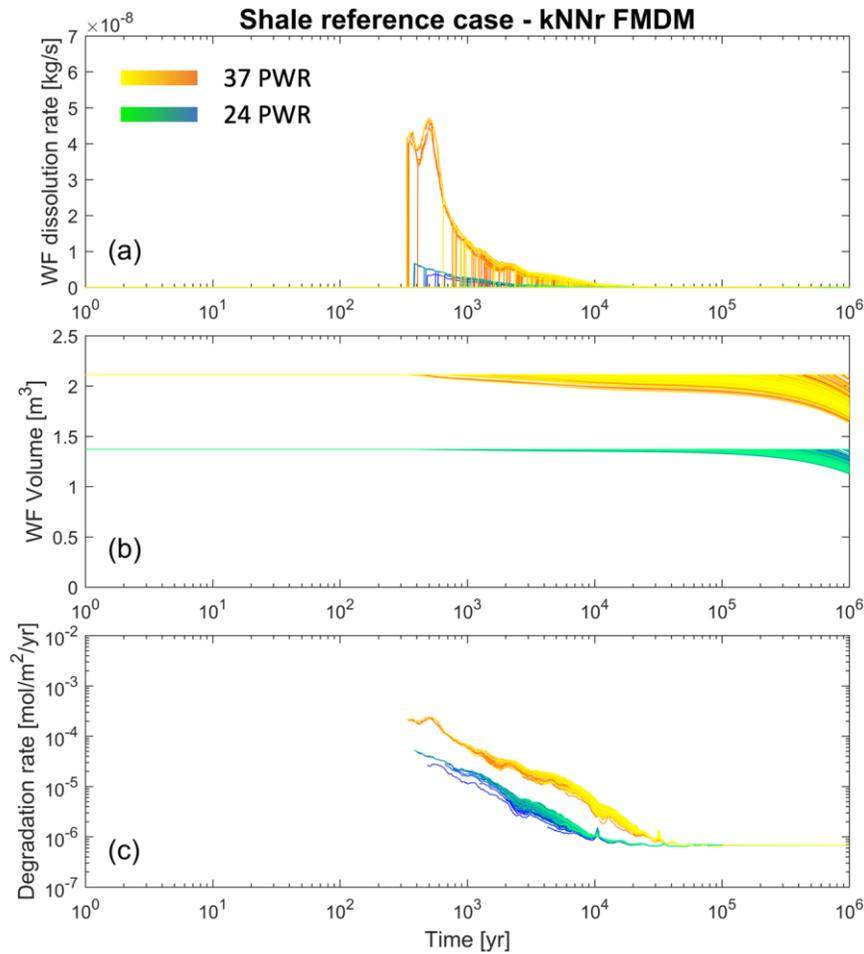


Fig. B-22 kNNr FMD surrogate model results using the 250k training input data: (a) fuel matrix degradation rate (kg/s), (b) volume of remaining waste form (m³), and (c) specific degradation rate (mol/m²/yr)

Speed Comparison

The shale repository system simulations ran on 1024 processors. Tab. B-9 shows the speed comparison of surrogate model simulations for the current report (FY 2021) and the previous report (Mariner et al. 2020).

Tab. B-9 Speed comparison of the shale repository system simulations

| | Time consumed [sec] | | | | |
|------------|---------------------|------------|------------|------------|------------|
| | FDR | ANN FMDM | | kNNr FMDM | |
| | | FY 2020 | FY 2021 | FY 2020 | FY 2021 |
| Flow | 2.6771E+04 | 2.6574E+04 | 3.0344E+04 | 2.6275E+04 | 3.1033E+04 |
| Transport | 1.7499E+04 | 2.7329E+04 | 2.9256E+04 | 2.7224E+04 | 2.9917E+04 |
| Waste Form | 1.5847E+03 | 2.8125E+02 | 1.1646E+03 | 3.3449E+02 | 1.1437E+03 |

The time spent in the WF calculations increased about threefold from FY 2020 due to the larger training data sets. However, the time needed was still significantly less than the FDR model, and extremely low compared to the time that would have been needed for the fully coupled FMD process model. Times spent on WF calculations relative to the total time spent on flow and transport increased from 0.5% to 1.8% for ANN and from 0.6% to 1.7% for kNNr. As indicated in the 52-WP problem of Mariner et al. (2020), if the shale case had been run using the coupled FMD process model, the simulation of the FMD process model would have required much more time than flow and transport.

CONCLUSIONS

The results of the coupled PFLOTTRAN simulations clearly show the potential of surrogate models to enable accounting for detailed Fuel Matrix Degradation dynamics while keeping the computational cost of reservoir simulations manageable. While this is a successful proof of concept, there is room for improvement in the surrogate accuracy.

The main enhancement over Mariner et al. (2020) in this report is the updated sampling strategy for the generation of training data, which includes data from roughly 8 times as many samples in the parameter space with down sampling in the time dimension for each run by a factor of about 10. This led to a more balanced data density across all dimensions in the sample space. Together with a narrowing of the features to just the 4 that have the most impact on the UO₂ Surface Flux prediction, this improved sampling approach led to an improvement of the kNNr MSE error over the test set by a factor of 2 compared to Mariner et al. (2020). Besides this factor of two improvement, the updated approach indicates that adding yet even more samples in the configuration space will continue to improve the accuracy of kNNr.

As the coupled simulation results show, kNNr stills struggles when the query points are close to the edge of the training space, while the ANN approach appears to be less sensitive to this.

Some other approaches that may primarily benefit the kNNr approach are listed below:

- **Manhattan distance metric:** As shown in Mariner et al. (2020), the Manhattan distance tends to give better accuracy than the Euclidean metric for the current datasets and could be added to the KDTree 2 Fortran implementation.
- **Enhanced data conditioning:** Rather than simple log₁₀ transformation of the training data, kNNr may benefit from additional normalization so that all features are weighted appropriately when selecting the nearest neighbors. In this context, it may be beneficial to emphasize some features over others if they are more predictive for the Quantity of Interest. Note that data conditioning has a similar effect as the choice of the distance metric used for selecting the nearest neighbors.
- **Online accuracy monitoring:** With kNNr, we know that the prediction accuracy will decline the further the nearest neighbors are from the query point. This could be used to assess during run-time whether the kNNr results can be trusted.

The Fuel Matrix Degradation (FMD) model calculates spent fuel degradation rates as a function of radiolysis, redox reactions, electrochemical reactions, alteration layer growth, and diffusion of reactants through the alteration layer. It is a complicated model requiring a large number of calculations and iterations at each time step. Because of this, repository simulations, which are already expensive, cannot directly include the FMD process model, especially when hundreds or thousands of WPs breach.

Mariner et al. (2020) showed the potential of surrogate models for coupled FMD – Reservoir simulations. It showed that an ANN surrogate and a k-Nearest Neighbors regressor (kNNr) surrogate can emulate the FMD process model with reasonable accuracy and with good computational efficiency, but it identified areas where the kNNr surrogate accuracy could get improved. The current report implemented a new and improved sampling strategy for generating training data that allows the kNNr surrogate to get much better accuracy.

REFERENCES

- Ben-David, S. and S. Shalev-Shwartz (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge, United Kingdom, Cambridge University Press.
- Berg, Timothy M., Kyung Chang, Rosemary Leone, Daniel T. Seidl, Paul Mariner, Bert Debusschere. (2021a, May 19). *Machine Learning Surrogates for the Fuel Matrix Degradation Model*. Spent Fuel and Waste Science and Technology Workshop (SFWST). Sandia National Laboratory Report SAND2021-6132 PE.
- Berg, Timothy M., Kyung Chang, Rosemary Leone, Daniel T. Seidl, Paul Mariner, Bert Debusschere. (2021b, March 4). *Surrogate Modeling of Spent Fuel Degradation for Repository Performance Assessment*. SIAM Conference on Computational Science and Engineering 2021 (CSE21). Sandia National Laboratory Report SAND2021-2451C.
- Buck, E., J. L. Jerden, Jr., W. L. Ebert and R. S. Wittmann (2013). *Coupling the Mixed Potential and Radiolysis Models for Used Fuel Degradation*. FCRD-UFD-2013-000290. Washington D.C., U.S. Department of Energy.
- Cormen, T. H., C. E. Leiserson, R. L. Rivest and C. Stein (2009). *Introduction to Algorithms*. 3rd ed. Cambridge, Massachusetts, The MIT Press.
- Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.
- Helton, J. C. and F. J. Davis (2003). "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems." *Reliability Engineering & System Safety* 81(1): 23-69.
- Jerden, J. L., Jr., K. E. Frey, J. M. Copple and W. L. Ebert (2014). *ANL Mixed Potential Model for Used Fuel Degradation: Application to Argillite and Crystalline Rock Environments*. FCRD-UFD-2014-000490. Washington D.C., U.S. Department of Energy.
- Jerden, J., G. Hammond, J. M. Copple, T. Cruse and W. Ebert (2015b). *Fuel Matrix Degradation Model: Integration with Performance Assessment and Canister Corrosion Model Development*. FCRD-UFD-2015-000550. Washington, DC, US Department of Energy.
- Jerden, J., J. M. Copple, K. E. Frey and W. Ebert (2015a). *Mixed Potential Model for Used Fuel Dissolution - Fortran Code*. O. o. U. N. F. Disposition. FCRD-UFD-2015-000159. Washington, DC, US Department of Energy.
- Jerden, J., T. Cruse, J. Fortner, K. Frey and W. Ebert (2012). *Used Fuel Degradation and Radionuclide Mobilization: Experimental Plan for Electrochemical Corrosion Studies*. M4FT-12AN0806011. Illinois, Argonne National Laboratory.
- Jerden, J., V. K. Gattu and W. Ebert (2017). *Progress Report on Development of the Spent Fuel Degradation and Waste Package Degradation Models and Model Integration*. SFWD-SFWST-2017-000091, SFWD-SFWST-2017-000095. Lemont, Illinois, Argonne National Laboratory.
- Jerden, J., V. K. Gattu and W. Ebert (2018). *Update on Validation and Incorporation of a New Steel Corrosion Module into Fuel Matrix Degradation Model*. M4SF-18AN010301017, M4SF-18AN010302016. Illinois, Argonne National Laboratory.
- Kennel, M. B. (2004). *KDTREE 2: Fortran 95 and C++ software to efficiently search for near neighbors in a multi-dimensional Euclidean space*. <https://arxiv.org/pdf/physics/0408067.pdf>, Institute for Nonlinear Science, University of California.
- King, F. and M. Kolar (1999). *Mathematical Implementation of the Mixed-Potential Model of Fuel Dissolution Model Version MPM-V1.0*. Report No. 06819-REP-01200-10005 R00, Ontario Hydro, Nuclear Waste Management Division
- King, F. and M. Kolar (2003). *The Mixed-Potential Model for UO₂ Dissolution MPM Versions V1.3 and V1.4*. Report No. 06819-REP-01200-10104 R00, Ontario Hydro, Nuclear Waste Management Division.
- Mariner, P.E., T.M. Berg, K.W. Chang, B.J. Debusschere, R.C. Leone, D.T. Seidl (2020). *Surrogate Model Development of Spent Fuel Degradation for Repository Performance Assessment*. M3SF-

- 20SN010304044. SAND2020-10797 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Mariner, P. E., D. T. Seidl, B. J. deBusschere, J. Vo, J. M. Frederick and L. P. Swiler (2019b). High Fidelity Surrogate Modeling of Fuel Dissolution for Probabilistic Assessment of Repository Performance. SAND2019-4151 C. International High-Level Radioactive Waste Conference, April 16, 2019, Knoxville, Tennessee.
- Mariner, P. E., E. R. Stein, S. D. Sevougian, L. J. Cunningham, J. M. Frederick, G. E. Hammond, T. S. Lowry, S. Jordan and E. Basurto (2018). Advances in Geologic Disposal Safety Assessment and an Unsaturated Alluvium Reference Case. SFWD-SFWST-2018-000509, SAND2018-11858 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Mariner, P. E., L. A. Connolly, L. J. Cunningham, B. J. DeBusschere, D. C. Dobson, J. M. Frederick, G. E. Hammond, S. H. Jordan, T. C. LaForce, M. A. Nole, H. D. Park, F. V. Perry, R. D. Rogers, D. T. Seidl, S. D. Sevougian, E. R. Stein, P. N. Swift, L. P. Swiler, J. Vo and M. G. Wallace (2019a). Progress in Deep Geologic Disposal Safety Assessment in the U.S. since 2010. M2SF-19SN010304041, SAND2019-12001 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Mariner, P. E., W. P. Gardner, G. E. Hammond, S. D. Sevougian and E. R. Stein (2015). Application of Generic Disposal System Models. FCRD-UFD-2015-000126, SAND2015-10037 R. Sandia National Laboratories, Albuquerque, New Mexico.
- Murphy, Kevin P. Machine learning: a probabilistic perspective. MIT press, 2012.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay (2011). "Scikit-learn: Machine Learning in Python." Journal of Machine Learning Research 12: 2825-2830.
- Radulescu, G. (2011). Radiation Transport Evaluations for Repository Science, Letter Report. ORNL/LTR-2011/294. Oak Ridge, Tennessee, Oak Ridge National Laboratory.
- Rasmussen, C. E. and C. K. I. Williams (2006). Gaussian Processes for Machine Learning., MIT Press.
- Santner, T., B. Williams and W. Notz (2003). The Design and Analysis of Computer Experiments. New York, New York, Springer.
- Sevougian, S. D., E. R. Stein, T. LaForce, F. V. Perry, T. S. Lowry, M. Nole and K. W. Chang (2019). GDSA Repository Systems Analysis FY19 Update. SAND2019-11942R. Sandia National Laboratories, Albuquerque, New Mexico.
- Simpson, T. W., V. Toropov, V. Balabanov and V. F.A.C. (2008). Design and analysis of computer experiments in multidisciplinary design optimization: A review of how far we have come or not. Proceedings of the 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Victoria, British Columbia, Canada. AIAA Paper 2008-5802.
- SKB (2006). Long-term safety for KBS-3 repositories at Forsmark and Laxemar – a first evaluation. SKB TR-06-09. Stockholm, Sweden, Svensk Kärnbränslehantering AB.
- Storlie, C. B., L. P. Swiler, J. C. Helton and C. J. Sallaberry (2009). "Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models." Reliability Engineering & System Safety 94(11): 1735-1763.
- Tieleman, Tijmen, and Geoffrey Hinton. "Lecture 6.5-rmsprop, coursera: Neural networks for machine learning." University of Toronto, Technical Report (2012).

**Appendix C. NEAR-TERM GDSA THRUSTS IN THE DISPOSAL
RESEARCH R&D 5-YEAR PLAN UPDATE**

**APPENDIX C. NEAR-TERM GDSA THRUSTS IN THE
DISPOSAL RESEARCH R&D 5-YEAR PLAN UPDATE**

Disposal Research 5-Year Plan Thrusts: Near Term FY21

| Thrust | Description | Tech. Area | Term |
|--|---|------------|------|
| Geologic Disposal Safety Assessment | | | |
| G01 | <p><i>Advanced simulation capability</i></p> <p>Near term advancements in PFLOTRAN capability will continue to focus on high priority topical areas identified in the 2019 Disposal Research R&D Roadmap Reassessment while building on previous accomplishments. These efforts will include continuing to advance high-temperature simulation capability, implementing material-specific waste package degradation models, and addressing the coupled thermo-hydro-mechanical (THM) processes affecting buffer evolution, and the coupled thermo-hydro-chemical (THC) processes affecting radionuclide transport. Multi-fidelity model implementation, including mechanistic models derived from detailed process understanding, reduced order models, and machine learning emulators, enhances computational efficiency and dovetails with integration of advanced UQ/SA methods.</p> <p>Other software development tasks will include ongoing development of open-source biosphere simulation software; addition of capability to dfnWorks (Hyman et al. 2015), software for generating discrete fracture networks and simulating particle transport; and release of an open source version of Vorocrust (Abdelkader et al. 2020), an automated meshing tool for generating conforming meshes of complex engineered and geologic features.</p> | GDSA | Near |
| G02 | <p><i>Uncertainty and sensitivity analysis</i></p> <p>Uncertainty quantification and sensitivity analysis (UQ/SA) methods in GDSA Framework will be advanced in two ways. First, the team will continue to identify and demonstrate methods consistent with the current standard of practice that add value to deep geologic repository performance assessment such as surrogate (meta) modeling, variance decomposition, multifidelity analysis, and evaluation of model form uncertainty. The team will continue to take a leadership role in international collaboration on these topics. Second, the team will begin to evaluate the reliability of methods dependent on surrogate models through techniques such as cross-validation and development of quantitative metrics for assessing goodness of surrogates.</p> | GDSA | Near |

| Thrust | Description | Tech. Area | Term |
|--|---|------------|------|
| Geologic Disposal Safety Assessment | | | |
| G03 | <p><i>Workflow</i></p> <p>Transparent, traceable workflows increase stakeholder confidence and user-friendliness. In the next 1 to 2 years, GDSA automation will be further developed using NGW and two additional important workflows will be established: an open source framework (scripted in Python) that automates software (PFLOTRAN) verification testing; and a workflow that streamlines data transfer from the geologic model to the meshing software and ultimately to the simulator.</p> <p>Integration of GDSA Framework with the Online Waste Library will be initiated. This integration will provide quality-assured radionuclide inventories for simulations involving defense-related waste streams.</p> | GDSA | Near |
| G04 | <p><i>Repository systems analysis</i></p> <p>In the next 1 to 2 years, a main priority will be simulation and analysis of the salt and crystalline reference cases developed for the DECOVALEX2023 task. This task will drive development of models of bentonite backfill evolution and waste package degradation (crystalline) and of salt consolidation and creep; and advance understanding of uncertainties associated with simulation and analysis methods. GDSA will continue to integrate with other technical areas to advance analyses of direct disposal of DPCs, understanding of high-temperature FEPs, and scenario development methodology.</p> | GDSA | Near |
| G05 | <p><i>Geologic modeling</i></p> <p>Geologic modeling involves two primary efforts: generation of representative 3-dimensional (3D) regional geology models that inform reference case concepts and simulations, and development of an interactive web-based application (https://gis1.inl.gov/regionalgeology/) for visualizing argillite, salt, and crystalline formations in the US. In the next 1 to 2 years, geologic models of a generic unsaturated alluvial basin and of regional argillite stratigraphy will be linked into the meshing workflow described above; and 3D subsurface visualization tools will be developed within the web application.</p> | GDSA | Near |

(This page is intentionally blank.)

Appendix D. NEAR-TERM GDSA THRUSTS AND THEIR MAPPED ROADMAP ACTIVITIES

APPENDIX D. NEAR-TERM GDSA THRUSTS AND THEIR MAPPED ROADMAP ACTIVITIES

5-Year Plan Thrusts and their Related Roadmap Activities

| Thrust | Act. ID | Activity | Tech. Area | Term |
|--|---|--|------------|------|
| Geologic Disposal Safety Assessment | | | | |
| G01 | <i>Advanced simulation capability</i> | | GDSA | Near |
| | P-06 | (Pseudo) colloid-facilitated transport model | | |
| | P-07 | Intrinsic colloid model | | |
| | P-08 | Other missing FEPs (processes) in PA-GDSA | | |
| | P-09 | Surface processes and features | | |
| | P-11 | Pitzer model | | |
| | P-12 | WP degradation model framework | | |
| | P-13 | Full representation of chemical processes in PA | | |
| | P-14 | Generic capability development for PFLOTRAN | | |
| | P-15 | Species and element properties | | |
| | P-16 | Solid solution model | | |
| | P-17 | Multi-component gas transport | | |
| G02 | <i>Uncertainty and sensitivity analysis</i> | | GDSA | Near |
| | P-05 | Disruptive events | | |
| | P-10 | Uncertainty and sensitivity analysis | | |
| G03 | <i>Workflow</i> | | GDSA | Near |
| | P-01 | CSNF repository argillite reference case | | |
| | P-02 | CSNF repository crystalline reference case | | |
| | P-03 | CSNF repository bedded salt reference case | | |
| | P-04 | CSNF repository unsaturated zone (alluvium) reference case | | |
| | P-10 | Uncertainty and sensitivity analysis | | |
| G04 | <i>Repository systems analysis</i> | | GDSA | Near |
| | P-01 | CSNF repository argillite reference case | | |
| | P-02 | CSNF repository crystalline reference case | | |
| | P-03 | CSNF repository bedded salt reference case | | |
| | P-04 | CSNF repository unsaturated zone (alluvium) reference case | | |
| | P-05 | Disruptive events | | |
| | P-08 | Other missing FEPs (processes) in PA-GDSA | | |
| | P-09 | Surface processes and features | | |
| | P-10 | Uncertainty and sensitivity analysis | | |
| | P-12 | WP degradation model framework | | |
| | P-13 | Full representation of chemical processes in PA | | |

| Thrust | Act. ID | Activity | Tech. Area | Term |
|--|--------------------------|--|------------|------|
| Geologic Disposal Safety Assessment | | | | |
| G05 | <i>Geologic modeling</i> | | GDSA | Near |
| | P-01 | CSNF repository argillite reference case | | |
| | P-02 | CSNF repository crystalline reference case | | |
| | P-03 | CSNF repository bedded salt reference case | | |
| | P-04 | CSNF repository unsaturated zone (alluvium) reference case | | |
| | P-09 | Surface processes and features | | |
| | P-13 | Full representation of chemical processes in PA | | |

(This page is intentionally blank.)